Proteolytic Queueing on Persister Cells

Alawiah H. Abualrahi[1], Heather S. Deter[1] and Nicholas C. Butzin[1]

***Short Abstract* — Persistence is a rare phenotypic switch that allows bacteria to be resilient against several factors, including antibiotics. Proteases are important elements for the formation of, and perhaps maintenance of, the persistence state. Here we explore the role of proteases in persister cells using queueing theory, the study of waiting lines, which has previously been applied to biological systems based on computational models and synthetic biology principles. We have applied proteolytic queueing to affect protease activity and found that proteolytic queues lead to an increase in the number of persister cells during stationary phase.**

***Keywords* —persistence, proteolytic queueing, proteases**

I. Full Abstract

Bacterial persistence is considered a major contributing factor if antibiotic treatments fail, particularly in chronic and reoccurring infections[1] and it has been identified in nearly all major bacterial pathogens[2]. Persisters are a subpopulation commonly found in bacteria that can precipitate antibiotic resistance[3]. Persister cells tolerate antibiotics without genetic mutations or spore formation[4]. Antibiotics kill vegetative cells (non-persisters), but persister cells survive because they have slow metabolic processes compared to vegetative cells, especially slower translation[5]. Persistence is affected by multiple systems, such as toxin-antitoxin (TA) and proteases activity. TA systems are often considered key regulators of bacterial persistence; when toxins in these systems are at a higher level than their cognate antitoxins the resulting free toxin activity triggers persistence[6]. The proportion of antitoxin-to-toxin can be regulated at both the degradation and production levels. Many antitoxins of the TA systems are degraded by the same proteases, and it is likely that this common factor causes some level of coordination[7].

Proteases play a fundamental role in persistence. However, gene knockouts and overexpression have not yet identified how cellular regulation is influenced by changes in protease activity. To examine the effects of protease activity, we used queueing theory, in which one type of customer competes for processing by servers, that has traditionally been applied to systems such as computer networks and call centers[8]. The biological queueing theory principally assumes that there are limited processing resources in a cell. Synthetic systems have been engineered to demonstrate that proteolytic queues form and can coordinate synthetic circuits when proteases are a limited resource[9]. Proteolytic queueing does not affect growth rates and provides a tunable method of changing degradation rates. We can synthetically cause a proteolytic queue by inducing expression of a fluorescent protein linked to amino acid tag, which targets the protein for degradation by a specific protease. We have found that the level of persister cell increases over 100-fold upon induction. This effect on persister levels only occurs when queue formation is maintained during antibiotic treatment (quantification of persister relies on killing non-persister cells). Furthermore, induction during antibiotic treatment alone increases persistence and therefore transcription and translation must be occurring during the antibiotic treatment to maintain the queue. These results underscore the importance of both transcription and translation for persister cell viability.

While many known persister related genes may be affected by the queue, TA systems are likely not responsible for queueing effects. If there was a queued antitoxin, then proteolytic queueing should result in higher free toxin and fewer persisters. We observed substantially higher persister levels, thus it seems unlikely that TA systems are directly responsible. We are currently screening for the protein(s) responsible for increased persistence in response to the queue.

II References

[1] Fisher, R. A.; Gollan, B.; Helaine, S., Persistent bacterial infections and persister cells. Nat Rev Microbiol 2017, 15 (8), 453-464.

[2] Van den Bergh, B.; Michiels, J. E.; Wenseleers, T.; Windels, E. M.; Boer, P. V.; Kestemont, D.; De Meester, L.; Verstrepen, K. J.; Verstraeten, N.; Fauvart, M.; Michiels, J., Frequency of antibiotic application drives rapid evolutionary adaptation of *Escherichia coli* persistence. Nat Microbiol 2016, 1, 16020.

[3] Michiels, J. E.; Van den Bergh, B.; Verstraeten, N.; Michiels, J., Molecular mechanisms and clinical implications of bacterial persistence. Drug Resist Updat 2016, 29, 76-89preparation," in preparation.

[4] Wood, T. K.; Knabel, S. J.; Kwan, B. W., Bacterial persister cell formation and dormancy. Appl Environ Microbiol 2013, 79 (23), 7116-21.

[5] Maisonneuve, E.; Gerdes, K., Molecular mechanisms underlying bacterial persisters. Cell 2014, 157 (3), 539-48

[6] Chan, W. T.; Espinosa, M.; Yeo, C. C., Keeping the Wolves at Bay: Antitoxins of Prokaryotic Type II Toxin-Antitoxin Systems. Front Mol Biosci 2016, 3, 9.

[7] Muthuramalingam, M.; White, J. C.; Bourne, C. R., Toxin-Antitoxin Modules Are Pliable Switches Activated by Multiple Protease Pathways. Toxins (Basel) 2016, 8 (7).

[8] Butzin, N. C.; Hochendoner, P.; Ogle, C. T.; Hill, P.; Mather, W. H., Marching along to an Offbeat Drum: Entrainment of Synthetic Gene Oscillators by a Noisy Stimulus. *ACS synthetic biology* **2016,** *5* (2), 146-53.

[9] Butzin, N. C. & Mather, W. H. Crosstalk between Diverse Synthetic Protein Degradation Tags in *Escherichia coli*. ACS Synth Biol 7, 54-62, doi:10.1021/acssynbio.7b00122 (2018).

1. Department of Biology and Microbiology, South Dakota State University, Brookings, South Dakota 57007, United States. E-mail: alawiah.abualrahi@Jacks.sdstate.edu, nicholas.butzin@sdstate.edu.

Kinetic transport in a whole-cell model of *E. coli*

Eran Agmon[1] and Markus Covert[1]

***Abstract*— To claim that we understand the behavior of a cell, we must be able to predict how would that cell would respond to changes in its environment. The transport system is a primary mediator of a cell's response to changes in external nutrient concentrations — transport fluxes can quickly drive changes to internal molecular pools. Over longer time scales, this can influence gene expression. As part of our effort to extend whole-cell modeling into more complex environments, the work described here integrates a kinetic model of amino acid transport with a whole-cell model of *Escherichia coli*.**

***Index Terms*— Whole-cell model, transport, amino acids, environment modeling, Escherichia coli**

I. Background

Whole-cell modeling aims to predict cellular behavior by integrating heterogeneous datasets and parameters into a detailed mechanistic model of cellular processes. In 2012 our lab reported the first whole cell model, of *Mycoplasma genitalium* [1]. More recent effort has focused on extending the framework to *Escherichia coli* [2]. In addition to having ten times more genes than *M. genitalium*, and 50-100 times as many molecules, *E. coli* can also grow in a wide variety of environmental conditions. This vastly increases the scope its modeling requirements.

Adding a kinetic model of transport is part of an effort to extend the whole-cell model of *E. coli* into more complex environments. Previous transport in whole-cell models was implemented as part of a metabolic sub-model, in a similar method to flux-balance analysis [3]. But in more complex environments, this approach is likely to produce inaccurate simulations. The current work introduces a transport sub-model that is deterministic with respect to environmental conditions, and constrains the whole-cell model's behavior.

II. Results

We identified 466 transport processes controlled by 239 unique transporters in *E. coli* [4]. These include diffusion through pores, electrochemical potential-driven transporters, and active transporters such as those driven by ATP. Amino acids alone are transported by 71 transport processes, controlled by 43 unique transporters. As a practice run for a more complete kinetic transport sub-model, we here focus only on these amino acid transport processes

Kinetic model of transport. Each transport process is defined as a rate law that determines flux between periplasmic and cytoplasmic molecular pools as a function of transporter and substrate concentrations. We use convenience kinetics; a general rate laws that can be expanded into any possible kinetic mechanism defined by a biochemical network [5].

Parameter estimation. Due to the difficulty of characterizing non-soluble, membrane-associated proteins, only few kinetic parameters for transport have been experimentally measured. Because of this, we have curated other measurements related to transport such as uptake rates, intra- and extra-cellular metabolite concentrations, and transporter concentrations. A genetic algorithm is used to search the parameter space of a given rate law for parameters values that best fit all of the experimental data. As needed, we simplify the rate laws — for example, by omitting concentration terms that are thought to be saturating.

Integration with whole-cell model. We test different implementation strategies to simulate how transport impacts metabolism. The whole-cell model without kinetic transport provides predictions of transport fluxes in a few known conditions – these become targets for the kinetic model parameterization. We then remove transport from the metabolic sub-model and instead use the determined uptake and secretion rates to calculate changes to internal metabolite pools. The metabolism sub-model responds to these changes by adjusting flux distributions that meet flexible objectives.

III. Conclusion

As whole-cell models continue to develop, they will need to predict behavior in increasingly complex environments. This will enable simulations to investigate cellular response to environments that are impossible to reproduce in the wet lab. As a first step towards this goal, we have here focused more closely on dynamic environments that vary only in few nutrient concentrations — in this case, amino acids. Getting transport right and seamlessly integrated will be an essential piece of predicting cellular response.

[1]Department of Bioengineering, Stanford University, Stanford, CA, 94305, USA. Email: eagmon@stanford.edu, mcovert@stanford.edu

References

[1] J. R. Karr, J. C. Sanghvi, D. N. Macklin, M. V. Gutschow, J. M. Jacobs, B. Bolival Jr, N. Assad-Garcia, J. I. Glass, and M. W. Covert, "A whole-cell computational model predicts phenotype from genotype," *Cell*, vol. 150, no. 2, pp. 389–401, 2012.

[2] D. N. Macklin, N. A. Ruggero, J. Carrera, H. Choi, T. A. Horst, J. C. Mason, M. M. DeFelice, I. Maayan, M. L. Paull, J. H. Morrison, E. Agmon, G. Sun, S. Akhter, S. R. Bray, D. S. Weaver, I. M. Keseler, P. D. Karp, and M. W. Covert, "Simultaneous cross-evaluation of heterogeneous *E. coli* datasets via mechanistic simulation," (in review).

[3] E. W. Birch, M. Udell, and M. W. Covert, "Incorporation of flexible objectives and time-linked simulation with flux balance analysis," *Journal of theoretical biology*, vol. 345, pp. 12–21, 2014.

[4] M. H. Saier Jr, C. V. Tran, and R. D. Barabote, "Tcdb: the transporter classification database for membrane transport protein analyses and information," *Nucleic acids research*, vol. 34, 2006.

[5] W. Liebermeister and E. Klipp, "Bringing metabolic networks to life: convenience rate law and thermodynamic constraints," *Theoretical Biology and Medical Modelling*, vol. 3, no. 1, p. 41, 2006.

Optimization of Transport to Multiple Protein Demand Sites in the Neuron

Anamika Agrawal[1] and Elena F Koslover[1]

***Abstract*— Extended neuronal cells require active transport to replenish proteins lost to turnover and decay throughout their periphery. We develop a quantitative model to explore how an active transport system can be optimized to deliver proteins to multiple targets along a neuronal projection, balancing broad vs rapid distribution of proteins manufactured in the cell body. Our model is applied specifically to the dynamics of protein exchange by fusion and fission between a dual population of stationary and motile mitochondria, quantifying the extent to which such a transport mechanism can support spatially heterogeneous metabolic demand.**

***Index Terms*— Intracellular Transport, Neuron, Mitochondria**

I. Introduction

Due to their large size, neuronal projections tend to have spatially heterogeneous needs for protein delivery from the cell body in order to maintain proteomic homeostasis [1]. This variability presents a challenge to intracellular transport systems that must deliver proteins simultaneously to multiple targets spaced throughout the cell periphery. For example, nodes of Ranvier situated at periodic intervals throughout the axon require a timely supply of ion channels, transporters and membrane proteins [2]. Similarly, synaptic bouton regions have high metabolic demands and thus require localized delivery of mitochondria and associated proteins [3] Long-range transport and delivery of proteins from the soma to peripheral targets occurs via active transport vehicles like the golgi-derived vesicles, mitochondria, and protein aggregates [4]. We use physical modeling to delineate how an active transport system could be optimized to efficiently deliver and replenish proteins at multiple sites along a neuronal projection. Specifically, we explore how the spatial distribution of metabolic proteins is regulated by dual populations of stationary and motile mitochondria that exchange contents through dynamic fission and fusion events.

II. Methods and Approach

We adopt analytical and computational methods to model protein transport along the axon. We assume that proteins are produced at a constant rate at the cell body at one terminal and are transported along a one-dimensional axonal projection. Subcompartments such as the nodes of ranvier and stationary mitochondria at synaptic boutons are treated as protein 'sinks' interspersed through the axonal domain. Protein carriers move processively through the domain, with enhanced probability of stopping and exchanging proteins at these localized regions of high demand. Different proteins are expected to have sinks with different spacing, as well as differing decay rates during transport and in the sink regions. A full exploration of the parameter regimes for this model will thus enable its application to multiple neuronal delivery systems.

Because targets are located sequentially along the neuronal projection, enhanced delivery of proteins at proximal locations necessarily decreases the flux to more peripheral targets. It is thus possible to identify an optimum set of transport parameters (velocity, stopping rates, and processivity) for efficient and robust delivery to multiple targets, depending on the rate of protein decay during transit and the target site density. While the analytical model is used to develop mean-field steady-state results, we leverage stochastic simulations to explore the effect of confining proteins within discrete carrier vesicles. In particular, we investigate the effect of mitochondrial fusion into clusters of varying sizes on the exchange of proteins through a dynamic system of motile and stationary mitochondria within neuronal axons. We also model the response rate for re-establishing proteomic homeostasis in the event of an injury or other local enhancement in protein turnover.

III. Conclusion

We develop a transport model to delineate the optimum parameters for protein delivery to multiple target sites along an extended domain. The model is applied to quantify the efficiency of protein distribution by a dynamic population of fusing and splitting mitochondria in neurons.

References

[1] V. H. Cornejo, A. Luarte, and A. Couve, "Global and local mechanisms sustain axonal proteostasis of transmembrane proteins," *Traffic*, vol. 18, no. 5, pp. 255–266, 2017.

[2] H. C. Lai and L. Y. Jan, "The distribution and targeting of neuronal voltage-gated ion channels," *Nature Reviews Neuroscience*, vol. 7, no. 7, p. 548, 2006.

[3] Q. Cai, M. L. Davis, and Z.-H. Sheng, "Regulation of axonal mitochondrial transport and its impact on synaptic transmission," *Neuroscience research*, vol. 70, no. 1, pp. 9–15, 2011.

[4] A. Brown, "Axonal transport of membranous and nonmembranous cargoes: a unified perspective," *The Journal of cell biology*, vol. 160, no. 6, pp. 817–821, 2003.

[1]Department of Physics, University of California, San Diego. Email: ana057@ucsd.edu

Designing protease-based enzymatic logic circuits

Deepak K. Agrawal[1], Elliott Dolan[2], Sagar D. Khare[2] and Eduardo D Sontag[3]

***Abstract*— We present mathematical models of protease enzyme-based biosensors that are capable of detecting two different inputs and produce as output either an OR or an XOR Boolean function. These biosensors employ post-translational modifications of synthetic proteins which are designed to interact in the presence of chemical and physical signals at timescales of seconds to minutes. To achieve desired functionality, we optimize the operation of each reaction network through rigorous numerical analysis, and the improvements are explained using sensitivity analysis. We demonstrate that each biosensor can meet the performance specifications at biologically plausible reaction parameters. These results suggest that our approach is viable for constructing fast reaction networks for biosensing applications with a capability of molecular computing.**

***Index Terms*— Logic gates, protease, enzyme, mathematical modeling.**

I. Purpose

Cellular signaling networks perform efficient computation in a complex environment by sensing and processing a multitude of chemical and physical signals into various responses that play a significant role in cell metabolism and function. This biological computation typically utilizes chemical reaction networks that follow defined input-output characteristics to generate a diversity of response outputs in response to a set of inputs. Inspired by these, several elementary and complex synthetic biological circuits have been developed to demonstrate proof-of-principle for biological computation [1]. These synthetic circuits typically rely on gene expression, which is used as an elementary signal and can take several hours to days to effectively propagate and respond to a stimulus. In recent times, an alternative paradigm has been developed that uses post-translational protein modifications or protein fragments mediated by enzymes to realize fast molecular computation [2], [3]. To further this approach, here we present mathematical models of protease-based elementary logic circuits for the rapid detection and response to chemical and optical inputs.

II. RESULTS

A. Biosensor components

In our design, currently under experimental implementation, each biosensor uses chemically-induced dimerized (CID) [4] and light-induced dimerized switches (LID) [5] where a small molecule or blue light acts to bring two proteins or protein fragments together respectively. This interaction allows inactive split protease enzymes which are fused to each dimerizer protein to come in close proximity, thus restoring activity of the protein by reconstituting it in response to the input stimulus. This process converts a chemical or physical input signal into a sufficient concentration of an active protease enzyme. Thus, each switch is capable of producing the same type of output in response to different kinds of inputs (small molecule or blue light). We develop ordinary differential equation (ODE) models and analyze the dynamic response of each switch through numerical and mathematical analysis under realistic reaction conditions.

B. Constructing biosensors with Boolean logic computation

CID and LID switches can be networked in order to make more complex circuits that can process versatile inputs on a fast timescale. As proof-of-principle, we construct the model of a biosensor capable of producing a high output (fluorescence) when either or both inputs are present, thereby demonstrating an OR gate functionality. Our approach is advantageous over others [2] in that, instead of designing a completely new reaction network for each logic gate, CID and LID switches can be considered as elementary building blocks for new logic gates. To demonstrate this capacity, we extend the OR gate design to achieve XOR gate functionality by allowing the two protease enzymes used by each switch to degrade each other. Therefore, when both the inputs are present, reconstituted proteases annihilate each other, resulting in a low output.

III. Conclusion

Our work serves as a guide to the design of protease-based enzymatic biosensors for molecular computation, in a mathematical framework to evaluate and optimize circuit performance generalizable to other molecular logic circuits.

Work partially supported by grants DARPA FA8650-18-1-7800 and NSF 1817936.

[1,3]Bioengineering and Electrical and Computer Engineering Departments, Northeastern University. Email: agwal.deepak@gmail.com

[2]Chemistry and Chemical Biology Department, Rutgers University. Email: khare@chem.rutgers.edu

[3]Laboratory of Systems Pharmacology, Program in Therapeutic Science, Harvard Medical School. E-mail: e.sontag@northeastern.edu

References

[1] N. Nandagopal and M. B. Elowitz, "Synthetic biology: integrated gene circuits," *Science*, vol. 333, no. 6047, pp. 1244–8, 2011.

[2] X. J. Gao, L. S. Chong, M. S. Kim, and M. B. Elowitz, "Programmable protein circuits in living cells," *Science*, vol. 361, no. 6408, pp. 1252–1258, 2018.

[3] T. Fink, J. Lonzaric, A. Praznik, T. Plaper, E. Merljak, K. Leben, N. Jerala, T. Lebar, Z. Strmšek, and F. Lapenta, "Design of fast proteolysis-based signaling and logic circuits in mammalian cells," *J Nature chemical biology*, vol. 15, no. 2, p. 115, 2019.

[4] S. Voss, L. Klewer, and Y. W. Wu, "Chemically induced dimerization: reversible and spatiotemporal control of protein function in cells," *Current Opinion in Chemical Biology*, vol. 28, pp. 194–201, 2015.

[5] B. Kim and M. Z. Lin, "Optobiology: optical control of biological processes via protein engineering," *Biochemical Society Transactions*, vol. 41, pp. 1183–1188, 2013.

Yeast as bioreactors: Understanding the role of vacuolar size in biochemicals concentration

Gabriela Alvarez-Azanedo, B.A., and YH Mark Chan, Ph.D.

Department of Biology, San Francisco State University, San Francisco, CA 94132

Eukaryotic cells contain internal compartments (organelles) which house molecules involved in specialized biochemical reactions. Thus, organelles can be used as bioreactors. For instance, vacuoles in *Saccharomyces cerevisiae* yeast are dynamic organelles that are capable of manufacturing precursor compounds of liquid fuels such as gasoline. However, it is unknown if vacuolar size plays a role in the concentration of the compounds that vacuoles store. Here, we examine whether vacuolar size limits the amount of a compound vacuoles can accumulate. Our compound of interest is p-Ribosylamino Imidazole (AIR), an intermediate molecule in the '*de novo*' Purine Biosynthesis (DNPB) pathway. AIR accumulates in the vacuoles of *S. cerevisiae* mutants defective in the *ADE2* gene (also involved in the DNPB pathway) when grown in adenine-deprived medium. To study the capacity of different vacuole sizes to store AIR, *ade2* mutants will be imaged via confocal, fluorescent microscopy. These images will reveal three-dimensional information of the vacuoles and AIR content from fluorescent signatures. The fluorescence intensities from the vacuolar membrane (GFP-tagged) and AIR (naturally red fluorescent when exposed to oxygen) will be analyzed using ImageJ and MatLab and used as a proxy to determine a correlation between vacuole size and amount of accumulated AIR. Our expected results include a direct correlation between vacuolar size and the amount of AIR accumulated. Understanding how vacuolar size affects the amount of a compound accumulated within the vacuoles could shed light on finding effective ways to enhance yeast vacuoles as bioreactors to manufacture [other] compounds of interest in larger scales.

An Auto-catalytic Cell Intercalation Mechanism to Understand Tissue Elongation during Morphogenesis

Samira Anbari[1], Javier Buceta[2]

***Abstract*— Axial elongation during the development of all metazoans body plans is a crucial process that is not fully understood. Here, we propose a mechanism based on an auto-catalytic feedback between cell mechanical properties and cell signaling to address the problem of tissue elongation. The proposing mechanism relies on a non-trivial combination of the differential adhesion hypothesis and the concept of positional information that modulates the former in a location-dependent manner. Our numerical simulations use a vertex model approach and we implement the French Flag model and Turing patterning as positional information mechanisms to illustrate our proposal.**

I. Introduction

During development, the initial spherical symmetry of the zygote undergoes complex changes in size and shape to form different tissues/organs and implement the body plan [1]. In that regard, axis elongation is a key morphogenetic geometric transformation that relies on the regulation of growth, division, and motility due to the interplay between signaling and cell mechanical properties. In this context, different mechanisms have been proposed and it has been shown that differential growth, oriented divisions, active migration, and cellular intercalation processes contribute to tissue elongation in various contexts [2], [3]. Yet, open questions remain. The limb bud is a model system in morphogenesis to understand patterning and the directed developmental expansion of a tissue. However, the reported features of the limb bud outgrowth reveal that this process is far from being understood and set an intriguing elongation conundrum [4]. Herein, we propose a general framework to understand the process of tissue elongation that relies on the interplay between cellular mechanical properties and signaling mechanisms to provide cell positional information within a primordium.

II. Results

A. Mechano-signaling Feedback: The Auto-Catalytic Cell Intercalation Model

Cell intercalation driven by the differential adhesion hypothesis (DAH) cannot sustain elongation. The reason is that the DAH hypothesis assigns distinct mechanical properties to cell populations based on inherited cellular identity. Here, instead, we propose a DAH intercalation mechanism where adhesion depends on identities that are provided, dynamically, by positional information as determined by signaling cues rather than inherited.

B. Illustrating Auto-Catalytic Cell Intercalation Model using French Flag model and Turing patterning.

To illustrate our model, that we refer to as Auto-catalytic Cell Intercalation, we perform numerical simulations of growing tissues by means of a vertex model approach [5], and two well-known positional information mechanisms: the French Flag model and Turing patterning. The Auto-catalytic Cell Intercalation mechanism induces a noticeable elongation in tissues patterned by Morphogen gradients. However, robust axis elongation in Turing patterned tissues relies on synergetic interactions of anisotropic growth and cell intercalation. In both cases, our results show that if cell-cell adhesion is regulated by the cellular environment in a location-dependent manner, then intercalation can be sustained in time and division orientation is shaped by cell stretching (as reported in different contexts). As a result, tissues and cells elongate along perpendicular directions as described in experiments.

III. Conclusion

We propose a novel mechanism that fills the gap of knowledge about how mechano-signaling feedback may drive directional cell activities to induce sustained intercalation processes and explain both the cellular growth/division behavior and the elongation of tissues.

[1]Department of Chemical and Biomolecular Engineering, Lehigh University, Bethlehem, PA 18015. Email: saa416@lehigh.edu

[2]Department of Bioengineering, Department of Chemical and Biomolecular Engineering, Lehigh University, Bethlehem, PA 18015 . Email: jbuceta@lehigh.edu

References

[1] Gilbert and L. (Ed.), *Metamorphosis: a problem in developmental biology*. Springer Science and Business Media., 2013.

[2] D. A. Ede and J. T. Law, "Computer simulation of vertebrate limb morphogenesis," *Nature*, vol. 221, no. 5177, pp. 244–8, Jan 1969.

[3] A. Zuniga, "Next generation limb development and evolution: old questions, new perspectives," *Development*, vol. 142, no. 22, pp. 3810–20, Nov 2015.

[4] B. Boehm, H. Westerberg, G. Lesnicar-Pucko, S. Raja, M. Rautschka, J. Cotterell, J. Swoger, and J. Sharpe, "The role of spatially controlled cell proliferation in limb bud morphogenesis," *PLoS Biol*, vol. 8, no. 7, p. e1000420, Jul 2010.

[5] R. Farhadifar, J.-C. Röper, B. Aigouy, S. Eaton, and F. Jülicher, "The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing," *Curr Biol*, vol. 17, no. 24, pp. 2095–104, Dec 2007.

Towards another narrative for T lymphocyte motility from alternative statistical analysis

Francois Asperti, Judy Cannon, Melanie Moses

***Short Abstract* — T lymphocyte motility has been shown to be linked to T lymphocyte activation and to a larger extent to the development of an adaptive immune response. The purpose of this study is to show how statistical analysis is linked to the details of the narrative of a scientific investigation and to how allowing a more complex statistical model to describe the datasets of T lymphocytes motility can lead to the definition of an alternative - and non obvious - narrative for the explanation of T lymphocyte motility itself and, to a larger extent, to an alternative explanation of the involvement of T lymphocyte motility in T lymphocyte-dependent adaptive immune response.**

***Keywords* — T lymphocytes, Cellular motility, Multiscale statistical analysis.**

The study of in vivo T lymphocyte motility has been of interest for immunologists in the last two decades as T lymphocytes present an active behaviour allowing these cells to interact with many different cells and structures in the near environment. In this context, many teams have started to try and quantify several aspects of T cell behaviour in the organs where T cells interact with dendritic cells. Motility parameters such as translational speed (velocity) and rotational speed (turning angle) are two typical parameters quantified ; the main goal being to link T cell motility to T cell ability to interact with the environment and, to a larger extent, to the ability of T cells to get activated and initiate an immune response.

In vivo T lymphocyte motility is studied by monitoring labelled T cell motility in lymph nodes (or other types of secondary lymphoid organs) using a confocal microscope. Several microscope fields from several organs from several mice, each field having multiple fluorescent lymphocytes are taken into account for a statistical analysis of the data.

The aim of such a statistical analysis is to compare the statistical parameters of a "*wild type"* population of T lymphocytes to a different population of T lymphocytes (treated or KO for a certain gene) in order to understand the mechanistic role of molecular pathways involved in T cell motility and the potential benefits for understanding T cell activation and the more general concept of T cell-dependent immunity.
Even though this way of reasoning seems fair, a justified comparison of different cellular populations (experimental condition) of lymphocytes requires that a thorough study of potential statistical heterogeneities be performed for each condition (separately), so that the potential differences uncovered from statistical comparison is related, not to a lack of explanation for the multiple scales of complexity and heterogeneities in the statistical structure of each dataset, and rather to the *mechanistic* (structural/molecular/genetic) differences between the two populations of lymphocytes compared with each other in a study.

In other words, cell motility and more specifically lymphocyte motility interpretation have been recently brought to another level by taking into account in vivo analysis of cellular behaviour.

In this context, assumptions about potential statistical heterogeneities in the data can be made instead of considering a single process (from a single underlying distribution) of T cell instantaneous velocity or T cell instantaneous turning angle among **individual cell tracks**, **individual microscopic field**, **individual lymph node** and **individual mouse**.

Current results obtained by re-analyzing T lymphocyte data motility in lymph nodes suggests that taking into account a more complex statistical structure, involving multiple scales and heterogeneities, can allow, in the same way, for the definition of a new and relevant narrative to explain T cell activation in the context of a lymph node complex T cell behaviours. Instead of considering a single statistical population of behavioural events (implying only one underlying mechanistic process) of T cell behaviour, multiple groups of behaviour can be statistically defined and compared, these behaviours being included and studied in the context of a theoretical framework developed in parallel.

Design Guidelines For Sequestration Feedback Control

Ania-Ariadna Baetica[1], Yoke Peng Leong[2] and Richard M. Murray[3]

***Abstract*— Integral control is a type of feedback commonly used in mechanical and electrical systems to achieve perfect adaptation to disturbances. A proposed implementation of integral control in synthetic biological systems uses the strong sequestration binding between two biochemical controller species for feedback control. The unbound amount of the two sequestering biochemical species represents the integral of the error between the current and the desired state of the system being controlled. However, in practice, implementing integral control using sequestration feedback has been challenging due to the controller molecules being degraded and diluted inside cells. Furthermore, integral control can only be achieved by sequestration feedback networks with binding rates that satisfy stability constraints. In our work, we give guidelines for guaranteeing the stability and good performance (small steady-state error) of sequestration feedback networks. Our guidelines provide simple tuning options to obtain a flexible and practical biological implementation of sequestration feedback control. Using tools and metrics from control theory, we pave the path for the systematic design of synthetic biological systems.**

***Index Terms*— feedback control, synthetic biology, integral control, sequestration feedback, controller design**

I. Introduction

An important goal of synthetic biology is to engineer reliable, robust, and well-performing systems from standardized biological parts that can easily be combined together [1], [2], [3], [4], [5]. Nonetheless, synthetic systems can lack robustness and be sensitive to their biological implementation [6], [7], [8]. Differences in their biological parts, different model organism implementations [9], or different experimental conditions can cause synthetic systems to cease to function properly [10], [11]. This limits their applicability, as well as the engineering of more complex synthetic systems.

The development of synthetic biological systems is limited by the lack of consistent functionality, performance, and robustness. These limitations are also present in mechanical and electrical engineering. However, tools and concepts have been developed to ameliorate them [12], [13], [14], [15]. For example, the engineering design cycle is a framework for the iterative design, build, testing, and learning of engineered systems [15]. The engineering design cycle can also be iteratively applied to synthetic systems to achieve the desired performance standards. In this work, we use engineering principles and tools to design of a class of well-performing, robust synthetic biological systems.

A widely used tool to improve the performance of mechanical and electrical systems is feedback control. Feedback control allows a system to take corrective action based on the measured differences between the current and desired performance [15]. The foremost benefit feedback control provides to biological, mechanical, and electrical systems is robustness to uncertainty. Should the system undergo a change such an external disturbance, the feedback controller ensures that the system retains good performance properties such as small steady-state error and fast response time by correcting for the change. Additionally, feedback control stabilizes an unstable process and accelerates a slow process. Yet, if poorly designed, feedback control can inadvertently amplify noise and exacerbate instability [16], [15].

Feedback is ubiquitous in natural biological systems, where it serves to regulate their behavior. Examples of feedback control found in natural biological systems include the regulation of body temperature [17], circadian rhythms [18], calcium [19], and glycolysis [20]. In this work [21], we explore the use of a feedback controller for synthetic biological systems that relies on the sequestration reaction of two biological species. This sequestration feedback controller is illustrated in Figure 1. Examples of synthetic systems that use sequestration feedback include the concentration tracker in [22], the two bacterial growth controllers in [23], and the gene expression controller in [24]. An example of a natural sequestration feedback system uses sigma factor σ^{70} and anti-sigma factor Rsd [25] for the two sequestering species.

The sequestration feedback introduced in Figure 1 has been a promising implementation of feedback control that can achieve perfect adaptation. Perfect adaptation (whose engineering counterpart is integral control) ensures adaption to disturbances with no steady-state error. A biological system with perfect adaptation displays excellent robustness and performance in terms of steady-state error. Perfect adaptation is a desirable property that can be found both in natural biological systems and in synthetic systems [28], [29], [19], [30], [16]. Briat *et al.* [31] studied stochastic sequestration feedback systems with no controller species degradation and demonstrated that they can achieve perfect adaptation under certain conditions of stability. Nevertheless, we and others

This work was sponsored by the Defense Advanced Research Projects Agency, Agreement HR0011-17-2-0008. The content of the information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

[1]Department of Biochemistry and Biophysics, University of California, San Francisco. Email: ania-ariadna.baetica@ucsf.edu

[2]Department of Control and Dynamical Systems, California Institute of Technology. Email: ypleong@caltech.edu

[3]Department of Control and Dynamical Systems, and Division of Biology and Biological Engineering, California Institute of Technology. E-mail: murray@cds.caltech.edu

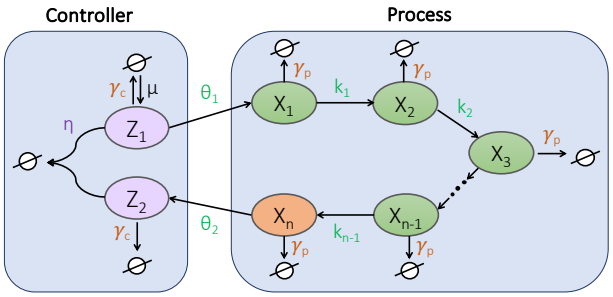


Fig. 1. **A class of sequestration feedback networks.** The sequestration feedback controller ensures that the process output species X_n (orange) tracks the reference signal set by the constitutive production of controller species Z_1 (purple) at rate μ. The controller species Z_1 and Z_2 (purple) bind together at rate η in a sequestration reaction to form an inactive complex represented by the empty set. The controller species Z_1 acts on the process input species X_1 at rate θ_1. The process output species X_n acts on the controller species Z_2 at rate θ_2. Process input species X_1 creates process species X_2 at rate k_1 and process species X_{n-1} creates process output species X_n at rate k_{n-1}. The controller and the process species are subjected to degradation and dilution, which is indicated by arrows pointing to empty sets. Rates γ_c and γ_p each encompass both degradation and dilution. For simplicity, we assume that the process species degradation rates γ_p (orange) are equal. Similarly, we assume that the controller species degradation rates γ_c (orange) are equal. An example process network that follows these assumptions implements the process species as proteins inside a bacterial cell. These simplifying assumptions match the setup in [26], [27].

have found the assumption of zero controller species degradation to be too restrictive for the practical implementation of sequestration feedback [32], [33], [5], [27].

Hence, current synthetic sequestration controllers may or may not be able to achieve perfect adaptation in a practical biological implementation. Therefore, we relax the requirement that the sequestration controller achieve perfect adaptation and we simulate and design sequestration feedback controllers with small steady-state error, large stability margin, and good disturbance rejection properties. These controllers are advantageous because they do not require an extremely precise implementation.

Previous research on sequestration feedback networks has determined conditions for their stability, along with their performance and robustness properties [31], [5], [26], [27]. Our current research [21] uses these properties of stability, performance, and robustness to find optimal designs for sequestration feedback networks. Additionally, we solve a case study design problem for a sequestration feedback network with two process species. Finally, we develop general guidelines for design of sequestration controllers for when the process network is already specified by challenging experimental constraints.

II. Conclusion

The development of a first generation of biological controllers for bacteria and yeast marks the beginning of an era when synthetic biological systems can function robustly and perform well. In this work [21], we have considered synthetic biological controllers implemented by a class of sequestration feedback networks. Using control theoretical methods, we have proposed designs that ensure stability, robustness, and good performance for these sequestration feedback network. We have offered tuning options for the strength of the controller species' sequestration reaction, as well as for the production and the degradation rates of both the process and the controller species. When possible, we have suggested biological parts for the practical implementations of these designs. Depending on the applications of interest to synthetic biology, we will benefit from multiple mechanisms for feedback control of synthetic systems and from multiple feedback controller designs.

References

[1] B. Canton, A. Labno, and D. Endy, "Refinement and standardization of synthetic biological parts and devices," *Nature Biotechnology*, vol. 26, no. 7, p. 787, 2008.

[2] A. Arkin, "Setting the standard in synthetic biology," *Nature Biotechnology*, vol. 26, no. 7, p. 771, 2008.

[3] R. Kwok, "Five hard truths for synthetic biology," *Nature News*, vol. 463, no. 7279, pp. 288–290, 2010.

[4] A. S. Khalil and J. J. Collins, "Synthetic biology: Applications come of age," *Nature Reviews Genetics*, vol. 11, no. 5, p. 367, 2010.

[5] Y. Qian and D. Del Vecchio, "Realizing 'integral control' in living cells: How to overcome leaky integration due to dilution?" *Journal of The Royal Society Interface*, vol. 15, no. 139, 2018.

[6] E. Yeung, A. Ng, J. Kim, Z. Z. Sun, and R. M. Murray, "Modeling the effects of compositional context on promoter activity in an e. coli extract based transcription-translation system," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*. IEEE, 2014, pp. 5405–5412.

[7] D. Del Vecchio, "Modularity, context-dependence, and insulation in engineered biological circuits," *Trends in biotechnology*, vol. 33, no. 2, pp. 111–119, 2015.

[8] E. Yeung, A. J. Dy, K. B. Martin, A. H. Ng, D. Del Vecchio, J. L. Beck, J. J. Collins, and R. M. Murray, "Biophysical constraints arising from compositional context in synthetic gene networks," *Cell Systems*, vol. 5, no. 1, pp. 11–24, 2017.

[9] S. Klumpp, Z. Zhang, and T. Hwa, "Growth rate-dependent global effects on gene expression in bacteria," *Cell*, vol. 139, no. 7, pp. 1366–1375, 2009.

[10] P. E. Purnick and R. Weiss, "The second wave of synthetic biology: From modules to systems," *Nature Reviews Molecular Cell Biology*, vol. 10, no. 6, p. 410, 2009.

[11] M. M. Gomez, R. M. Murray, and M. R. Bennett, "The effects of time-varying temperature on delays in genetic networks," *SIAM Journal on Applied Dynamical Systems*, vol. 15, no. 3, pp. 1734–1752, 2016.

[12] W. S. Levine, *The Control Systems Handbook: Control System Advanced Methods*. CRC press, 2010.

[13] B. Friedland, *Control system design: an introduction to state-space methods*. Courier Corporation, 2012.

[14] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback control theory*. Courier Corporation, 2013.

[15] K. J. Aström and R. M. Murray, *Feedback Systems: An Introduction For Scientists And Engineers*. Princeton University Press, 2008.

[16] D. Del Vecchio and R. M. Murray, *Biomolecular feedback systems*. Princeton University Press, 2015.

[17] J. Werner, "System properties, feedback control and effector coordination of human temperature regulation," *European Journal of Applied Physiology*, vol. 109, no. 1, pp. 13–25, 2010.

[18] M. J. Rust, J. S. Markson, W. S. Lane, D. S. Fisher, and E. K. O'Shea, "Ordered phosphorylation governs oscillation of a three-protein circadian clock," *Science*, vol. 318, no. 5851, pp. 809–812, 2007.

[19] H. El-Samad, J. Goff, and M. Khammash, "Calcium homeostasis and parturient hypocalcemia: An integral feedback perspective," *Journal of Theoretical Biology*, vol. 214, no. 1, pp. 17–29, 2002.

[20] F. A. Chandra, G. Buzi, and J. C. Doyle, "Glycolytic oscillations and limits on robust efficiency," *science*, vol. 333, no. 6039, pp. 187–192, 2011.

[21] Ania-Ariadna Baetica, Y. P. Leong, N. Olsman, and R. M. Murray, "Design guidelines for sequestration feedback networks," *bioRxiv*, 2018. [Online]. Available: https://www.biorxiv.org/content/early/2018/10/30/455493

[22] V. Hsiao, E. L. de los Santos, W. R. Whitaker, J. E. Dueber, and R. M. Murray, "Design and implementation of a biomolecular concentration tracker," *ACS synthetic biology*, vol. 4, no. 2, pp. 150–161, 2014.

[23] R. D. McCardell, S. Huang, L. N. Green, and R. M. Murray, "Control of bacterial population density with population feedback and molecular sequestration," *bioRxiv*, p. 225045, 2017.

[24] F. Annunziata, A. Matyjaszkiewicz, G. Fiore, C. S. Grierson, L. Marucci, M. di Bernardo, and N. J. Savery, "An orthogonal multi-input integration system to control gene expression in escherichia coli," *ACS Synthetic Biology*, vol. 6, no. 10, pp. 1816–1824, 2017.

[25] M. Jishage and A. Ishihama, "Transcriptional organization and in vivo role of the escherichia coli rsd gene, encoding the regulator of rna polymerase sigma d," *Journal of Bacteriology*, vol. 181, no. 12, pp. 3768–3776, 1999.

[26] Ania-Ariadna Baetica, "Design, analysis, and computational methods for engineering synthetic biological networks," Ph.D. dissertation, California Institute of Technology, 2018.

[27] N. Olsman, Ania-Ariadna Baetica, F. Xiao, Y. P. Leong, J. Doyle, and R. Murray, "Hard limits and performance tradeoffs in a class of sequestration feedback systems," *bioRxiv*, 2018. [Online]. Available: https://www.biorxiv.org/content/early/2018/09/26/222042

[28] J. E. Ferrell Jr, "Perfect and near-perfect adaptation in cell signaling," *Cell Systems*, vol. 2, no. 2, pp. 62–67, 2016.

[29] L. Goentoro, O. Shoval, M. W. Kirschner, and U. Alon, "The incoherent feedforward loop can provide fold-change detection in gene regulation," *Molecular Cell*, vol. 36, no. 5, pp. 894–899, 2009.

[30] T.-M. Yi, Y. Huang, M. I. Simon, and J. Doyle, "Robust perfect adaptation in bacterial chemotaxis through integral feedback control," *Proceedings of the National Academy of Sciences*, vol. 97, no. 9, pp. 4649–4653, 2000.

[31] C. Briat, A. Gupta, and M. Khammash, "Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks," *Cell Systems*, vol. 2, no. 1, pp. 15 – 26, 2016.

[32] J. Ang, S. Bagh, B. P. Ingalls, and D. R. McMillen, "Considerations for using integral feedback control to construct a perfectly adapting synthetic gene network," *Journal of theoretical biology*, vol. 266, no. 4, pp. 723–738, 2010.

[33] X. Ren, Ania-Ariadna Baetica, A. Swaminathan, and R. M. Murray, "Population regulation in microbial consortia using dual feedback control," in *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on*. IEEE, 2017, pp. 5341–5347.

Specificity in Protein-Protein Interaction Networks

A. Jane Bardwell[1] and Lee Bardwell[1]

Short Abstract — **Protein overexpression can lead to non-specific binding events that deplete a cell of functional complexes, and increase toxic aggregates. We have developed a theoretical framework for specificity in protein-protein interaction networks. We find that interaction specificity displays an unexpected relationship to protein concentrations. For example, reducing the concentration of a given protein will always improve the specificities of its interactions, but this will, of necessity, decrease the specificity of other protein-protein interactions in the cell. Thus, cellular concentrations (but not Kd's) are Pareto optimal with respect to specificities. Other intriguing findings emerging from this analysis will also be discussed.**

Keywords — **Specificity, protein-protein interaction networks, binding, theoretical analysis, evolution, cancer, neurodegenative disease.**

I. Purpose

Many human diseases and disorders, including cancer, birth defects, and neurodegeneration, are caused or exacerbated by genetic or epigenetic changes that result in protein overexpression or underexpression. The binding of any two molecules depends not only on their intrinsic affinity, but also on their concentration. As a consequence of the law of mass action, any two molecules will associate if their concentrations are high enough. For proteins that tend to be sticky, these "dangerous concentrations" can readily be reached in a cell, especially if the proteins are overexpressed. Indeed, multiple studies that have shown that promiscuous protein-protein interactions exist, can have deleterious consequences for cells and organisms, and can impose constraints on evolutionary change (1-6).

Intrinsic protein disorder and linear motif content have been found to be some of the best predictors of dosage sensitivity (7). Intrinsically-disordered regions of proteins are thought to be promiscuous and non-specifically sticky (8,9). Moreover, interactions mediated by linear motifs (such at the PXXP motif recognized by SH3 domains) are notoriously non-specific (10). Thus, strong evidence suggests that a propensity for concentration-dependent interaction promiscuity can lead to off-target, non-cognate interactions that have detrimental consequences for fitness and contribute to disease progression. Although there is an extensive literature on aggregation of monomers (11,12), there is little theoretical analysis of the consequences of protein expression changes on specificity in protein interactions networks.

II. Summary of Results

Building upon our previous work (13-15), we have recently developed a theoretical framework for the analysis of specificity in protein-protein interaction networks. Specificity results from the ability of a protein to preferentially recognize its own binding partner(s) while not binding to highly similar proteins in other pathways. This is determined both by the structure of the network (the wiring and the binding affinities) and the concentration of the component proteins. Increasing the concentration of any protein always decreases the specificity of its interactions – and this propagates through the network, affecting other interaction specificities. Indeed, concentrations are Pareto optimal with respect to specificity- changing a concentration to make one interaction more specific will make others less specific. Strong selection against off-target interactions will favor a maximin strategy where concentrations are set so as to maximize the minimum specificity. Furthermore, for high specificity, we find that concentrations should be below Kds, thus providing a rationalization for the modest affinity of docking interactions mediated by short linear motifs.

III. Conclusion

These new findings provide insight into the structure of, and selective pressures upon, protein interaction networks.

References

1. McConkey, E. H. (1982). PNAS U S A 79, 3236-3240.
2. Levy, E. D., De, S., and Teichmann, S. A. (2012). PNAS USA 109, 20461-20466.
3. Geiler-Samerotte, K. A., et al (2011). PNAS U S A 108, 680-685.
4. Oromendia, A. B., Dodgson, S. E., and Amon, A. (2012) Genes Dev 26, 2696-2708.
5. Yang, J. R., et al (2012) PNAS USA 109, E831-840.
6. Sopko, R., et al. (2006). Mol Cell 21, 319-330.
7. Vavouri, T., et al. (2009). Cell 138, 198-208.
8. Cumberworth, A., et al (2013). Biochem J 454, 361-369.
9. Uversky, V. N., Oldfield, C. J., and Dunker, A. K. (2008). Annu Rev Biophys 37, 215-246.
10. Tompa, P., Davey, N. E., Gibson, T. J., and Babu, M. M. (2014). Mol Cell 55, 161-169.
11. Veitia, R. A., Bottani, S., and Birchler, J. (2008) Trends Genet 24, 390.
12. Cohen, S. I., Vendruscolo, M., Dobson, C. M., and Knowles, T. P. (2012). J Mol Biol 421, 160-171.
13. Haney, S., Bardwell, L., and Nie, Q. (2010). BMC Syst Biol 4, 119.
14. Komarova, N. L., Zou, X., Nie, Q., and Bardwell, L. (2005). Mol Syst Biol 1, 2005 0023.
15. Bardwell, L., Zou, X., Nie, Q., and Komarova, N. L. (2007). Biophys J 92, 3425-3441.

[1]Department of Developmental & Cell Biology, Center for Complex Biological Systems, University of California, Irvine. E-mail: bardwell@uci.edu

Authors:

Effie E Bastounis, Julie A Theriot

Title:

Mechanical forces govern interactions of *Listeria monocytogenes* with host epithelial cells

Abstract:

The epithelium of the gut is the first barrier that must be circumvented by the intracellular bacterial pathogen *Listeria monocytogenes* (LM) during infection to initiate systemic spread. After LM establishes infection of an epithelial cell, it can perform cell-to-cell spread, using a form of actin-based motility to move through the cytoplasm, form a protrusion at the host cell surface that pushes into a recipient cell, and then becomes engulfed by the recipient cell before escaping into its cytosol to repeat this process. During cell-to-cell spread, LM engages host intercellular junctions and potentially reprograms host cells in ways that are still not yet fully understood.

We hypothesized that bacterial infection of host epithelial cells with LM would elicit changes in the biomechanics of infected hosts, including changes in the cell-extracellular matrix forces they exert while residing on deformable matrices mimicking their natural environment. To test our hypothesis, we performed traction force microscopy on MDCK epithelial cell monolayers infected with low levels of LM and measured the forces that cells exerted on the soft matrices throughout the course of several days post-infection. We found that infected host cells weaken the forces they exert on their substratum significantly over the first few days of infection as compared to nearby uninfected cells or compared to control uninfected cells. By monitoring the movement of host cells' nuclei, we also found that uninfected cells move towards the center of the infection focus, squeezing the infected cells and initiating the formation of highly packed domains of softer infected cells as revealed through atomic force microscopy. Over the course of 2 days post-infection, these domains start attaining a 3D "mound" configuration and squeezed infected cells get extruded. The orientation of cells around infection foci becomes highly anisotropic and uninfected cells orient themselves with their major axis perpendicular to the tangent of the focus. Through immunofluorescence microscopy and Western blotting, we have set out to find candidates that might be responsible for these changes and we have strong evidence that host innate immunity pathways are likely involved in that process.

Our work has aimed to identify what biomechanical changes occur during long-term infection of epithelial cells with LM. The weakening of cell-extracellular matrix forces of infected cells and the creation of extrusion domains underline the dynamic remodeling capability of epithelial tissue and might hint on a potential mechanism employed by host epithelial cells to contribute to clearance of bacterial infection.

A Novel Fluorescent Protein Reporter for Size of Live Human Cells

Daniel F. Berenson[1], Evgeny Zatulovskiy[1], Shicong Xie[1], and Jan M. Skotheim[1]

***Short Abstract* — Cell size is a critical but understudied determinant of cellular physiology. Research has been hindered by inadequate tools for measuring size in living cells. We applied a novel strategy to measure the size of single human cells using a nuclear-localized fluorescent protein expressed from a constitutive promoter. We validated this method by comparing it to established alternatives, and found that our fluorescence measurements are more robust and less dependent on image segmentation than the commonly-used measurement of nuclear volume. Finally, we examined how cell size and concentrations of key regulatory proteins together control the G1 to S phase transition.**

***Keywords* — Cell size, cell cycle, protein concentrations, live cell imaging.**

I. Introduction

CELL size influences a variety of biological processes, including biosynthesis and mitochondrial efficiency [1]. Moreover, cell size sets the denominator for concentration measurements used in mass-action kinetics reaction rates in ordinary differential equations mathematical models. Despite this importance, variation in cell size has received comparatively little quantitative research attention, perhaps because it has been relatively difficult to accurately measure the sizes of single living cells.

A number of techniques exist to measure cell size, but none is perfectly suited for all experiments. Coulter counters and flow cytometers provide snapshots of cell populations, but cannot follow individual cells through time; confocal microscopy and quantitative phase microscopy are difficult to combine with time lapse imaging; and fluorescence exclusion and microchannel resonance require custom equipment [2-3].

We hypothesized that we could exploit the fact that most proteins in a cell are expressed at a constant concentration – that is, their amount is proportional to the cell size – to accurately measure cell size by measuring a single fluorescent protein species.

II. Results

We used the constitutively expressed *EF1α* promoter to drive expression of mCherry-NLS in human mammary epithelial cells, then compared fluorescence to several other methods of measuring cell size. We used flow cytometry and widefield imaging to show that mCherry-NLS amount is proportional to forward scatter, a total protein dye, and nuclear volume.

A. Sensitivity to image segmentation

Measuring nuclear volume is another convenient and commonly used technique for estimating cell size, but also requires image segmentation to distinguish foreground nuclei [4]. We found that mCherry measurements were more robust than nuclear volume to changing the image segmentation threshold. In live time lapse images, we also found that mCherry was substantially less noisy than nuclear volume.

B. Analysis of G1 size control

We applied our fluorescence measurements of cell size to understand how cells control their size at the G1/S transition. Using live cell imaging and the FUCCI cell cycle phase marker, we observed that small-born cells spend longer in G1 than large born-cells, enabling them to catch up in size, while there is no such size-dependent effect on S-G2 length. We were also able to measure the decreasing concentration of the cell cycle inhibitor Rb as cells grow during G1, which provides a mechanism linking cell growth to cell cycle progression at the G1/S transition [5].

III. Conclusion

A constitutively expressed fluorescent protein offers a convenient and straightforward approach for studies into the effects of cell size on physiology [6].

References

[1] Miettinen TP, Björklund M (2016) Cellular allometry of mitochondrial functionality establishes the optimal cell size. *Dev Cell* 39.

[2] Cadart C, et al. (2018) Size control in mammalian cells involves modulation of both growth rate and cell cycle duration. *Nat Comm* 9, 3275.

[3] Son et al. (2015) Resonant microchannel volume and mass measurements show that suspended cells swell during mitosis. *J Cell Biol* 211, 757-763.

[4] Liu S et al. (2018) Size uniformity of animal cells is actively maintained by a p38-MAPK-dependent regulation of G1-length. *Elife* 7, e26947.

[5] Zatulovskiy E, Berenson DF, Topacio BR, Skotheim JM (2018) Cell growth dilutes the cell cycle inhibitor Rb to trigger cell division. *Biorxiv* 470013.

[6] Berenson DF, Zatulovskiy E, Xie S, Skotheim JM "Constitutive expression of a fluorescent protein reports the size of live human cells," in preparation.

[1]Department of Biology, Stanford University, California. E-mail: berenson@stanford.edu; skotheim@stanford.edu.

Acknowledgements: This work was funded by NIH grant R01 GM115479 and T32 GM007365.

Deciphering the code of cytokines with φ-evo

François Bourassa[1], Sooraj Achar[2], Angela Lee[2], Emanuel Salazar-Cavazos[2], Paul François[1]*, and Grégoire Altan-Bonnet[2]*

***Abstract*— T cells secrete and detect cytokines to communicate during a collective immune response, giving rise to rise to complex dynamical behaviour. Our experiments suggest that time series of cytokines encode the quality (ability to trigger a response) and the quantity of peptides (antigens) presented to T cells. How can T cells decode this information? We have used φ-evo, an *in silico* evolution algorithm, to find simple biochemical networks that could predict peptide quality based on cytokine time series. We found that a simple network computing the product of two cytokines can achieve a robust classification of peptide quality.**

***Index Terms*— cytokines, in silico evolution, T cell, biochemical network, collective computation**

I. Background information

TCells can be activated by antigens (peptides) binding to their antigen receptors, but they also need other stimulatory signals to sustain the activation on longer time scales [1]. Cytokines are an important part of these signals. They are messenger proteins that T cells can both produce and detect. They establish a global communication between immune cells by diffusing through tissues [2], to ensure the coherence of their collective response.

Research has shown that cytokines have multiple functions in the immune system. For instance, IL-2 stimulates T cell proliferation, controls their differentiation [3], increases expression of IL-2 receptors, but inhibits its own production, leading to counterintuitive scaling laws in the immune response [4]. Moreover, cytokines interact with one another: for example, IL-2 competes with other cytokines for receptor components, sometimes causing a given cytokine to have opposite effects [3]. The importance of such interactions demands that many cytokines be investigated jointly; yet, the majority of studies have examined one cytokine at a time. In total, there are around 30 well-documented cytokines in the human body, but we are still lacking quantitative principles to explain systematically their intricate mechanisms [5].

Our research tries to answer this need for a quantitative study of the dynamics of multiple cytokines together. Advances in this direction would notably be of interest for the field of cancer immunotherapy.

[1]Department of Physics, McGill University, Montreal, QC, Canada. Email: paul.francois2@mcgill.ca, francois.bourassa4@mcgill.ca

[2]Immunodynamics Section, Cancer and Inflammation Program, Center for Cancer Research, National Cancer Institute, Bethesda, MD, USA. Email: gregoire.altan-bonnet@nih.gov

*To whom correspondence should be addressed.

II. Research summary

A. Hypothesis and methods

Our goal is to model the dynamics of multiple cytokines during an immune response by T cells. We have performed experiments in which we measured the time course of the concentration of seven cytokines after exposing T cells to an antigenic peptide, for different peptides, at various concentrations. Our measurements indicate that different peptides trigger the secretion of distinct mixtures of cytokines. We have therefore conjectured that cytokine levels encode peptide quality (ability to stimulate T cell activation) and quantity, and we wanted to decipher this code.

To achieve this, we used the φ-evo software [6], an algorithm that simulates darwinian evolution of biochemical networks selected for a given function. We used this algorithm to evolve small biochemical networks that take cytokine concentrations as an input, and output antigen quality or quantity. Such networks can shed light both on the logic of the cytokine encoding and on possible mechanisms by which T cells could *decode* that information.

B. Results

Up to now, we have focused on predicting antigen quality. We have found a simple phosphorylation network that computes the product of the logarithm of IL-2 and either IL-6 or TNF-α. It can classify peptides in three quality categories and has been cross-validated across nine experimental data sets. The classification exhibits temporal dependence; it improves over time. This suggests decoding mechanisms such as a gene that would be activated jointly by pSTAT3 and pSTAT5 (proteins downstream of IL-2 and IL-6 signalling).

References

[1] J. M. Marchingo *et al.*, "Antigen affinity, costimulation, and cytokine inputs sum linearly to amplify T cell expansion," *Science*, vol. 346, no. 6213, p. 1123, Nov. 2014.

[2] G. Perona-Wright, K. Mohrs, and M. Mohrs, "Sustained signaling by canonical helper T cell cytokines throughout the reactive lymph node," *Nature Immunology*, vol. 11, no. 6, pp. 520–526, Jun. 2010.

[3] R. Spolski, P. Li, and W. J. Leonard, "Biology and regulation of IL-2: from molecular mechanisms to human therapy," *Nature Reviews Immunology*, Aug. 2018.

[4] K. E. Tkach *et al.*, "T cells translate individual, quantal activation into collective, analog cytokine responses via time-integrated feedbacks," *eLife*, vol. 3, p. e01944, 2014.

[5] G. Altan-Bonnet and R. Mukherjee, "Cytokine-mediated communication: a quantitative appraisal of immune complexity," *Nature Reviews Immunology*, Feb. 2019.

[6] A. Henry, M. Hemery, and P. François, "φ-evo: A program to evolve phenotypic models of biological networks," *PLOS Computational Biology*, vol. 14, no. 6, p. e1006244, Jun. 2018.

A Model to Study Distribution of Fitness Effects in Genetic Networks

R G Brajesh[1], Supreet Saini[1]

Short Abstract — **Mutations give rise to genetic variation resulting in an increase in the fitness of the population. A number of experimental, computational and theoretical studies report distribution of beneficial and deleterious mutations in organisms/systems. However, a precise, quantitative distribution of mutations is not understood. In this study, we consider the lactose utilization system in *E. coli* and define its fitness with the help of cost-benefit analysis. Using this, we develop a computational framework to employ mutations in the system and evaluate distributions linked to beneficial or deleterious mutations.**

***Keywords* — Distribution of fitness effects, *E. coli*, lactose utilization system, microbial evolution.**

I. PURPOSE

RANDOM mutations can be beneficial, deleterious or neutral in nature for an evolving species. The frequency distribution associated with new mutations is called as distribution of fitness effects (DFE). Obtaining DFE has been a challenge mainly because of its experimental limitations [1]. Limited availability of experimental data about point mutations even for a smaller organism like virus also makes the development of theoretical model difficult [2]. In this study, we chose lactose utilization system in *E. coli* and develop a computational framework to evaluate distribution of fitness effects[3]. Our results highlight that beneficial mutations are distributed exponentially but frequency distribution of deleterious mutations is far more complex. In addition, these distributions does not change qualitatively because of the starting fitness or the precise location in the fitness landscape. With the help of our computational framework, we also explore the epistatic interactions between beneficial mutations and quantitatively show that the spread of effect of beneficial mutation is drastically different in different genetic background.

II. RESULTS

A. Distribution of fitness effects

We use a simple mathematical model to obtain dynamics of gene expression in lactose utilization system and with the help of cost-benefit framework obtain fitness of the organism. Model parameter corresponding to particular fitness value was obtained, then distinct point mutations was introduced, and fitness was evaluated. This process was repeated for 10,000 times to obtain distribution of fitness effects.

We found that DFE of beneficial mutations was found to be best represented by an exponential distribution when the system is at low initial fitness ($0.001f_{max}$). The deleterious mutations can also be represented by an exponential distribution in this case. This simulation was tested for 100 different parameters and we see that the distribution remain same qualitatively.

B. Effect of starting fitness on DFE

When the previously obtained simulation was repeated for different starting fitness ($0.01f_{max}$, $0.1f_{max}$ and $0.5f_{max}$), we note that, DFE associated with beneficial mutation in all three cases can be represented by exponential distribution. However, the fitting of exponential distribution becomes less accurate with increase in initial fitness. On the other hand, DFE of deleterious mutations associated with all three fitness levels cannot be represented any standard distribution.

C. Epistatic interactions between beneficial mutations

When beneficial mutations were introduced in different beneficial background we note that resultant effect of the mutation can be quite drastically different in different parameter background.

III. CONCLUSION

Overall, we provide a computational framework to explore the distribution of beneficial and deleterious mutations and the dynamics of beneficial mutations in a controlled laboratory setup.

REFERENCES

[1] H. A. Orr, "The distribution of fitness effects among beneficial mutations," *Genetics,* vol. 163, Apr 2003.

[2] R. Sanjuan, A. Moya, and S. F. Elena, "The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus," *Proc Natl Acad Sci U S A,* , Jun 01 2004.

[3] E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. Van Oudenaarden, "Multistability in the lactose utilization network of Escherichia coli," *Nature,* vol. 427, pp. 737-40, Feb 19 2004.

Acknowledgements: This work was funded by Department of Science and Technology (Grant Number: SB/S3/CE/060/2015).

[1]Department of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai, India, 400076. Email: saini@che.iitb.ac.in

Bayesian detection of diffusive heterogeneity

Julie A. Cass[1], C. David Williams[1] and Julie A. Theriot[1,2]

***Abstract*— Cells are crowded and spatially heterogeneous, complicating the transport of organelles, proteins and other substrates. The diffusion constant partially characterizes dynamics in complex cellular environments but, when taken on a per-cell basis, fails to capture spatial dependence of diffusivity. Measuring spatial dependence of diffusivity is challenging because temporally and spatially finite observations offer limited information about a spatially varying stochastic process. We present a Bayesian framework that estimates diffusion constants from single particle trajectories, and predicts our ability to distinguish differences in diffusion constants, conditional on how much they differ and the amount of data collected.**

***Index Terms*— diffusion, intracellular transport, single particle tracking, Bayesian statistics**

I. Introduction

Diffusion is essential for the intracellular transport of organelles, proteins and substrates, and is commonly characterized through analyses of single particle tracking (SPT) in live-cell images [1], [2], [3]. While powerful analyses from SPT have indicated the complexity of transport in live cells, the spatial variation of the diffusion constant remains poorly characterized. This can be attributed to challenges in disentangling effects of biological heterogeneity and limited sampling of a stochastic process [4], [5]. To address these challenges, we developed a Bayesian framework to estimate a posterior distribution of the possible diffusion constants underlying SPT dynamics. This framework can be used to generate a look-up table predicting the detectability of differences in diffusion constants, conditional on the ratio of their values and amount of trajectory data collected.

II. Methods

We simulate particle diffusion in a range of homogeneous diffusion constants, and digest the resulting trajectories into frame-to-frame displacements. Using an inverse-gamma conjugate prior, we make the conservative guess that any order of magnitude of diffusion constant is equally likely. The set of displacements in a single trajectory are used to generate a posterior inverse-gamma distribution estimating the probability that any given diffusion constant was used to generate the trajectory. This distribution peaks near the true diffusion constant and has a width corresponding to the confidence interval of our estimate, which is largely determined by the trajectory length. Given a pair of posteriors derived for trajectories with differing underlying diffusion constants, we can characterize their similarity by computing the Kullback-Leibler divergence. This metric acts as a single-value estimation of how well we can analytically distinguish that trajectories were generated from different diffusion constants. For longer trajectory lengths, stochastic variations will be less dominant, increasing distinguishability.

III. Preliminary results

To assess the conditional feasibility of computationally detecting differences in diffusivity, we generate a landscape of the KL divergence between posteriors generated from pairs of simulations, with varying trajectory lengths and differences in diffusivity . To further correct for stochastic variations in simulations, the KL divergence reported for each entry in the landscape is the mean value from thousands of replicates. We find that, using this method, diffusivities differing by a factor of 1.5 or more can be easily distinguished when at least 50 timepoints are reported for each trajectory. This landscape offers a look-up table for estimating the number of frames that must be acquired experimentally to distinguish diffusivities to a desired precision. This framework could therefore play a valuable role in describing the feasibility of and requirements for experiments addressing the spatial heterogeneity of the intracellular diffusive environment.

IV. Conclusion

The spatial heterogeneity of diffusion may have major impacts in the transport of essential cellular substrates but remains largely uncharacterized. To shed light on the feasibility of resolving spatial from stochastic drivers of diffusive heterogeneity in trajectory data, we developed a framework for predicting our ability yo detect differences in diffusivity, conditional on the amount of experimental data collected. Our framework can therefore be used to inform the design of experiments aimed to characterize the spatial dependence of diffusivity across cells.

References

[1] G. M. Lee, A. Ishihara, and K. Jacobson, "Direct observation of brownian motion of lipids in a membrane," *PNAS*, vol. 88(14), pp. 6274–6278, 1991.

[2] M. J. Saxton and K. Jacobson, "Single-particle tracking: Applications to membrane dynamics," *Annu. Rev. Biophys. Biomol. Struct.*, vol. 26, pp. 373–399, 1997.

[3] S. C. Weber, A. J. Spakowitz, and J. A. Theriot, "Bacterial chromosomal loci move subdiffusively through a viscoelastic cytoplasm," *PRL*, vol. 104(23), no. 238102, 2010.

[4] M. T. Valentine, P. D. Kaplan, D. Thota, J. C. Crocker, T. Gisler, R. K. Prud'homme, M. Beck, and D. A. Weitz, "Investigating the microenvironments of inhomogeneous soft materials with multiple particle tracking," *Phys Rev E*, vol. 64, no. 061506, 2001.

[5] T. J. Lampo, S. Stylianidou, B. M. P, P. A. Wiggins, and A. J. Spakowitz, "Cytoplasmic RNA-protein particles exhibit non-gaussian subdiffusive behavior." *Biophysical Journal*, vol. 112(3), pp. 532–542, 2017.

[1]Allen Institute for Cell Science, Seattle WA USA.
[2]Department of Biology, University of Washington, Seattle WA USA.

Temporary oncogene inhibition leads to irreversible cell cycle arrest

Jeremy B. Chang[1], Danielle L. Swaney[1], Erica K. Stevenson[1], Lani F. Wu[2], Steven J. Altschuler[2], and Nevan J. Krogan[1]

***Short Abstract* — Typical dosing schedules for targeted cancer therapies aim to deliver as much drug as can be tolerated. We investigated the response of the HER2-overexpressing breast cancer cell line SK-BR-3 to various dosing schedules of the targeted therapy lapatinib, a specific HER2 inhibitor. Surprisingly, doses as short as an hour induced a sustained cell cycle arrest. We characterized the molecular underpinnings of this arrest using mass spectrometry and found that cells appear to be in a pre-restriction point-like state in which both the SCF and APC/C are inactive. We present a quantitative model of this irreversible cell fate decision.**

***Keywords* — HER2, cell cycle arrest, lapatinib, irreversible.**

I. Background

THE Human Epidermal Growth Factor Receptor (HER) family controls cell proliferation, and members of the family are genetically altered in nearly a fifth of all cancers [1]. Cell culture models of HER family-driven cancers commit apoptosis in response to HER family inhibitors. However, typically only a limited fraction of cells commit apoptosis, and these inhibitors have limited clinical efficacy.

Typical dosing schedules attempt to deliver as much drug as possible for as long as can be tolerated by the patient. However, evidence in the literature suggests in certain cases, such as in some leukemias, short treatment durations are sufficient to have clinical impact [2]. Other studies have highlighted the importance of scheduling and dynamics in cancer therapies [3,4]. We investigated the effect of treatment frequency and duration of the HER2 inhibitor lapatinib on the HER2-overexpressing cell line SK-Br-3.

II. Results

We used a combination of quantitative fluorescence video microscopy, mass spectrometry, and mathematical modeling to understand how treatment dynamics affects cellular fate.

A. Short treatment durations induce sustained cell cycle arrest.

We found that lapatinib treatment durations as short as one hour were sufficient to induce a sustained (>5 days) cell cycle arrest in SK-Br-3 cells. We treated cells with varying durations of 2 μM lapatinib. The medium was refreshed at a maximum of every two days. Cells were monitored using an widefield epifluorescence microscope for five days. By quantifying the cell number over time, we found that cells treated with longer than an hour stopped dividing.

B. Mass spectrometry and microscopy of arrest reveals that cells are in a pre-restriction point-like state.

We analyzed untreated and drug-treated cells by Data Independent Analysis Mass Spectrometry (DIA-MS) and Phosphorylation Mass Spectrometry (phospho-MS). We found that cells displayed a proteome and phosphoproteome indicative of arrest prior to the restriction point.

Furthermore, by using fluorescence ubiquitylation-based cell cycle (FUCCI) sensors [5,6], we determined that both of the major cell cycle-related ubiquitin ligases, the SCF and APC/C ubiquitin ligases, were inactive.

C. An ordinary differential equations (ODE)-based model suggests a mechanism for maintaining cell cycle arrest.

We created a simple ODE model of the relevant cell cycle regulators based on our mass spectrometry and microscopy data. Our model highlights the importance of negative feedback in maintaining the cell cycle arrest.

III. Conclusion

We have discovered and characterized a novel cell cycle arrest in response to a short pulse of HER2 inhibition in cancer cells. Inducing such cell cycle arrest may prove to useful clinically for limiting the growth of cancer. We are validating our findings in other cancer cell lines treated with other targeted therapies.

References

[1] cBioPortal for Cancer Genomics. (2019). Available at: https://www.cbioportal.org/.

[2] Shah, N. P. *et al.* Transient Potent BCR-ABL Inhibition Is Sufficient to Commit Chronic Myeloid Leukemia Cells Irreversibly to Apoptosis. *Cancer Cell* **14,** 485–493 (2008).

[3] Lee, M. *et al.* Sequential Application of Anticancer Drugs Enhances Cell Death by Rewiring Apoptotic Signaling Networks. *Cell* **149,** 780–94 (2012).

[4] Chen, S.-H. H., Forrester, W. & Lahav, G. Schedule-dependent interaction between anticancer treatments. *Science (New York, N.Y.)* **351,** 1204–8 (2016).

[5] Sakaue-Sawano, A. *et al.* Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression. *Cell* **132,** (2007).

[6] Bajar, B. T. *et al.* Fluorescent indicators for simultaneous reporting of all four cell cycle phases. *Nature Methods* **13,** 993–996 (2016).

[1]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, USA.

[2]Department of Pharmaceutical Sciences and Pharmacogenomics, University of California, San Francisco, USA.

Optimal response to pathogen evolution in immune repertoires

V. Chardès[1], T. Mora[1†] and A. Walczak[1†]

***Abstract*— In order to target threatening pathogens, the adaptive immune system performs a continuous reorganisation of its immune cell repertoire constrained on the antigenic load. This emergent self-organised repertoire results from a variety of biological processes such as affinity maturation or thymic selection. Previous work [1], [2], formulated a theory predicting immune repertoire organisation, assuming that it results from an optimization problem with biological constraints. Here we extend this formalism to a co-evolutionary setting, where we explicitely consider the effects of a continuously mutating viral population and immune cell maturation processes on the dynamics of immune repertoire adaptation.**

I. Motivations

Viruses and immune cells of the adaptive immune system engage in a continuous co-evolutionary process, as escaping viral mutations and strong immune response exert a feedback on one another. Previous work [1] provided an effective optimization framework predicting the optimal immune repertoire organisation for a stationary antigenic population. Further, describing immune adaptation to time dependent pathogen statistics as a sequential Bayesian estimation, [2] recovered biologically relevant behaviors such as memory production and attrition. However, on the side of viral evolution, these previous approaches were unable to adress a mutating viral population constantly creating unknown antigenic phenotypes. On the side of the immune system, even though the optimal immune organisation obtained was proven to be reached through a receptor competition for antigens, no constraints such as the need to go through cell maturation process have been imposed.

II. Effective framework

The considered optimality principle [1] states that the immune system minimizes a cost based on its average belief of each viral strain fraction, this average belief depicting the immune system's confidence about the existence of a viral strain. This minimization process provides the optimal distribution of immune receptors corresponding to the immune-estimated antigenic landscape. For a time-dependent viral population, the immune system is described as a sequential Bayesian estimator, continuously performing two steps (Fig. 1) : a prediction of the future viral population between encounters, based on the immune system prior belief of the viral dynamics, and a Bayesian update of this prediction each time a new viral strain is encountered by the immune system.

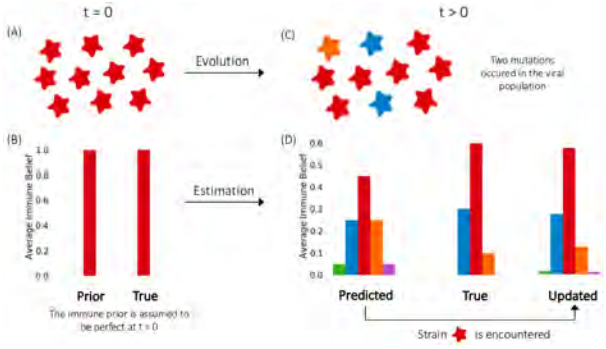


Fig. 1. Sketch of the immune system adaptation as a sequential Bayesian estimation. (A) and (B), the initial viral population is uniform, and the immune system has a prior matching the its true composition. (C) and (D), as soon as a viral strain is encountered, the immune system corrects its prediction through sequential Bayesian update. This updated belief now plays the role of immune prior in (B).

III. Results

We consider the evolution of viral strains under a Wright-Fisher process with antigen dependent mutations, taken as mutations onto nearest neighbours in the antigenic space. In this non-stationary system we show that the average immune belief obeys a diffusion equation on the antigenic space, with no dependence on the genetic drift. This result highlights that the immune repertoire dynamics between two viral encounters only depends on mutations. On the contrary, we show that the genetic drift is critical in the sequential Bayesian update step, i.e. when the immune system reorganises its repertoire in response to an antigen recognition event. When mutations are rare compared to genetic drift, the immune system should allocate all its resources around the last met viral strains, while when mutations are more frequent the immune system should give less weight to new viral encounters, as the number of existing mutants increases. Adressing this range of mutation rates allows us explore the immune adaptation to a wide variety of realistic viral evolutionary regimes.

This work was funded by ERC N°724208 Struggle.

[1]Laboratoire de Physique, Ecole Normale Superieure, 75005 Paris, France chardes@lpt.ens.fr

[†]Equal contribution

References

[1] A. Mayer, V. Balasubramanian, T. Mora, and A. M. Walczak, "How a well-adapted immune system is organized," *Proceedings of the National Academy of Sciences*, vol. 112, no. 19, pp. 5950–5955, May 2015.

[2] A. Mayer, V. Balasubramanian, A. M. Walczak, and T. Mora, "How a well-adapting immune system remembers," *arXiv:1806.05753 [q-bio]*, Jun. 2018.

Disrupting Transcriptional Feedback Yields an Escape-Resistant Antiviral

Authors: Sonali Chaturvedi[1] , Marie Wolf[1] , Noam Vardi[1] , Matilda F. Chan[2,3] , Leor S. Weinberger[1,4,5]

Short abstract- **Drug resistance is a substantial clinical problem, with combination therapies often the only recourse. Here, we propose a novel antiviral approach that disrupts viral auto-regulatory circuits, which limits resistance by requiring multiple viral mutations. We provide proof-of-concept that DNA-based circuit-disruptor oligonucleotide therapies (C-DOTs) interfere with transcriptional negative feedback in human herpesviruses (CMV and HSV-1) thereby increasing viral transcription factors to cytotoxic levels. C-DOTs reduce viral replication >100-fold, are effective in high-viremic conditions where existing antivirals are ineffective, and show efficacy in mice. Strikingly, no C-DOT-resistant mutants evolved in >60 days of culture, in contrast to approved herpesvirus antivirals where resistance rapidly evolved. Oligonucleotide therapies that target feedback circuits could mimic combination therapy and represent escape-resistant interventions with broad applicability to viruses, microbes, and neoplastic cells.**

Keywords- **herpesvirus; autoregulatory circuit; transcriptional feedback; viral evolution; oligonucleotide therapy**

PURPOSE

From bacteria to cancers, it has long been recognized that drug-resistant mutants emerge quickly, causing significant morbidity and mortality[1-3]. Antiviral resistance in herpesviruses is of particular concern[4,5] with herpes simplex virus type 1 (HSV-1)—a leading cause of blindness—and human herpesvirus 5, cytomegalovirus (CMV)—a leading cause of birth defects and transplant failure—exhibiting substantial resistance to standard-of-care antivirals in the clinic[6,7]. Combination therapies, which limit resistance by necessitating multiple viral mutations, can be effective but increase the risk of off-target effects and associated toxicity and are absent for most viral diseases. Here, we present proof-of-concept for a novel approach that disrupts viral auto-regulatory circuits with a single molecule and limits resistance by requiring multiple viral mutations. We develop DNA-based circuit-disruptor oligonucleotide therapies (C-DOTs) that exploit this mechanism by interfering with transcriptional negative feedback in human herpesviruses (CMV and HSV-1) thereby increasing viral transcription factors to cytotoxic levels. C-DOTs reduce viral replication >100-fold, prevent emergence of resistant mutants in continuous culture, are effective in high-viremic conditions where existing antivirals are ineffective, and show efficacy in mice. Strikingly, no C-DOT-resistant mutants evolved in >60 days of culture, in contrast to approved herpesvirus antivirals where resistance rapidly evolved. Overall, the results demonstrate that oligonucleotide therapies targeting feedback circuits are escape resistant and could have broad therapeutic applicability to viruses, microbes, and neoplastic cells.

REFERENCES

1 Meylan, S., Andrews, I. W. & Collins, J. J. Targeting Antibiotic Tolerance, Pathogen by Pathogen. Cell 172, 1228-1238, doi:10.1016/j.cell.2018.01.037 (2018).

2 Lee, H. H., Molla, M. N., Cantor, C. R. & Collins, J. J. Bacterial charity work leads to population-wide resistance. Nature 467, 82-85, doi:10.1038/nature09354 (2010).

3 Goldberg, D. E., Siliciano, R. F. & Jacobs, W. R., Jr. Outwitting evolution: fighting drug-resistant TB, malaria, and HIV. Cell 148, 1271-1283, doi:10.1016/j.cell.2012.02.021 S0092-8674(12)00221-8 [pii] (2012).

4 Piret, J. & Boivin, G. Antiviral drug resistance in herpesviruses other than cytomegalovirus. Rev Med Virol 24, 186-218, doi:10.1002/rmv.1787 (2014).

5 Lurain, N. S. & Chou, S. Antiviral drug resistance of human cytomegalovirus. Clin Microbiol Rev 23, 689-712, doi:10.1128/CMR.00009-10 23/4/689 [pii] (2010).

6 Frobert, E. et al. Resistance of herpes simplex viruses to acyclovir: an update from a ten-year survey in France. Antiviral Res 111, 36-41, doi:10.1016/j.antiviral.2014.08.013 (2014).

7 Emery, V. C. & Griffiths, P. D. Prediction of cytomegalovirus load and resistance patterns after Ativiral chemotherapy. Proc Natl Acad Sci U S A 97, 8039-8044, doi:10.1073/pnas.140123497 (2000).

Affiliations: [1] Gladstone|UCSF Center for Cell Circuitry, Gladstone Institutes, San Francisco, CA 94158; [2] Francis I. Proctor Foundation, [3] Department of Ophthalmology, [4] Department of Biochemistry and Biophysics, and [5] Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94158

Angeline Chemel

The Relationship between Vacuole Inheritance and Biogenesis

Saccharomyces *cerevisiae*, commonly known as budding yeast, forms a single bud and divides asymmetrically during its cell cycle. Previous research has shown vacuoles are crucial for cell cycle progression and suggests that the mammalian lysosome (the equivalent of the yeast vacuole) may also play a crucial role in mammals. Inheritance is a key pathway for ensuring proper distribution of organelles from the mother yeast to its bud and understanding this highly regulated process may shed light onto cancer and tumor prevention. Although inheritance is a highly regulated process, it is unclear in what pattern inheritance occurs. Biogenesis is measured using genetically modified yeast with GFP bound to VPH1, a protein abundant in vacuole membrane. Inheritance is measured using FM-4-64, a red dye that's trafficked to the vacuole. The cells are followed through a full stage of cell division and are documented with 3-D images using a spinning disk confocal microscope. I hypothesized that inheritance occurs in a discrete pattern over time, however data is suggesting that inheritance may occur in a steady progressive pattern. Additionally, three mutant strains, Apl5Δ, Atg18Δ, and Vps11Δ, with varying vacuole phenotypes such as reduced vacuole size, enlarged vacuole size, and fragmented vacuoles, respectively, were observed. Data collected from these mutant strains are suggesting that vacuole morphology may affect the pattern of vacuole inheritance. In addition, preliminary biogenesis data is suggesting that the mother goes through a significant amount of vacuole biogenesis, as hypothesized, however it seems that the bud goes through zero biogenesis. This suggests that the bud relies completely on the mother to receive the proper amount of vacuoles.

Dynamics of Blood Flow in Liver Discriminate Hepatic Fibrosis

Hao Chen[1], Zhengyan Li[2], Lihong Ye[2], Yi Jiang[1]

Short Abstract — Cirrhosis affects liver functions, and is a significant public health problem. Presently only early stages of liver fibrosis are treatable, but they are asymptomatic and difficult to diagnose. The mechanism of fibrosis changing the mechanical properties of the liver tissue and altering the dynamic of blood flow is still unclear. In collaboration with clinicians specialized in hepatic fibrosis, we have developed a mechanical model to integrate our empirical understanding of fibrosis progression, which correlates the stages of fibrosis to mechanical properties of liver tissue and the blood flow changes. These results are supported by ultrasound Doppler measurements from hepatic fibrosis patients. These results promise a new noninvasive diagnostic tool for early fibrosis.

***Keywords* — liver, cirrhosis, hepatic fibrosis, mechanical, spring, collagen, stiffness, vein, fibrosis grading.**

I. PURPOSE

Cirrhosis is a condition in which scar tissue replaces healthy tissue and liver loses function, due to long-term damage. Its most common causes are excessive alcohol consumption, liver disease and hepatitis B&S[1]

Progression of hepatic fibrosis is strongly correlated with increasing rigidity of the liver tissue [2]. We hypothesize that, as tissue rigidity can affect the dynamics of blood flow, fibrosis stages may be detected non-invasively using blood flow measurements. To test this hypothesis, we built a computational model of blood flow in hepatic fibrosis as a function of fibrosis progression. We then analyzed clinical liver biopsy images and the corresponding ultrasound Doppler images to confirm the relationship between the fibrosis grade and blood flow dynamics.

II. METHOD

We modeled toxin distribution in the liver lobules that leads to tissue damage, activation of myofibroblast and deposition of collagen fibers. According to the understanding of liver progression, we used a fiber-pixel model to simulate the progression of liber fibrosis, where fiber and portal vein are constructed as image pixel. This simulation, as a function of increasing fibrosis grade, shows excessive collagen form polygonal patterns, resembling those found in pathology images.

We developed a simple harmonic oscillator to model blood flow as a function of tissue stiffness around the blood vessels. The calculations show that the fibrosis induced stiffness change can cause quantifiable changes in blood flow in the hepatic veins. We have analyzed biopsy images and ultrasound Doppler images from 200 patients with fibrosis ranging from grades 1 to 4. (**Figure 1)**

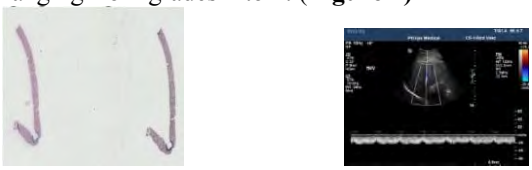


Figure 1: Left, biopsy image of liver; Right: ultrasound Doppler images.

III. Result

After analyzing the real data, this behavior follows the result of simulation, where the mechanical properties of blood flow were found to be correlated the extend of fibrosis. The clinical result is consistent with our assumption, the higher stage of hepatic fibrosis, the feedback value of wave pattern in the Doppler image. **(Figure 2)**

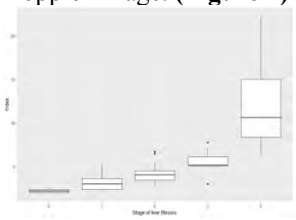


Figure 2. Grading liver fibrosis by blood flow

We have developed a preliminary model system, which suggests that the blood flow pattern can be used to determine the mechanical properties of the liver tissue, and the results have been shown to be consistent with clinical measurements. Based on this initial finding, we hypothesize that the increased tissue stiffness in the liver during the progression of liver fibrosis modulate the blood flow patterns in the portal vein and hepatic veins systematically, which can be quantified and detected noninvasively to accurately determine the stages of fibrosis, even in early stages.

Acknowledgements:

[1]Department of Mathematics and Statistics, Georgia State University, USA

[2] Shijiazhuang Fifth Hospital, China

REFERENCES

[1] Schuppand D, Afdhal NH (2008) Liver Cirrhosis, PMC

[2] Xu J (2001b) Amyloid-β peptides are cytotoxic to oligodendrocytes. J Neurosci **21**(RC11

Understanding Cell Size Homeostasis and Phenotypic Diversity during Bacteria Filamentation

Yanyan Chen[1], Javier Buceta[1,2]

***Abstract* — Under a number of conditions, bacteria develop filamentous morphologies. Yet, the principles driving their size homeostasis and phenotypic diversity remain unknown. Here, we present stochastic models supported by experiments to gain insight into these questions quantitatively. On the one hand, our results unravel the independence of size homeostasis on the division pattern and make compatible the *sizer* and *adder rules.* On the other hand, our preliminary results about phenotypic switching suggest that processes where both daughter cells are filamentous are less favorable.**

***Keywords* — Filamentous bacteria, Homeostasis, Size control, Phenotypic diversity, Stochastic modeling**

I. Background

UNder conditions like high temperature or pressure, antibiotic treatment, or genetic modifications, cells exhibit filamentous morphologies [1, 2]. Yet, filamentous cells are usually not considered in single-cell studies (likely because their morphology hinders the possibility of a robust statistical analysis).

In normal-sized strains, experiments spanning different species reveal various relations between division and born size. The relations were found to follow either one or a combination of two size control rules. In general, there are three rules – *timer*, *sizer,* and *adder*. The *timer* one suggests cell division after growing for a specific amount of time; *sizer* suggests cell division after reaching the critical division size; finally *adder,* which has been mainly discovered in *E.coli*, *B.subtilis* and *P.aeruginosa*, suggests a constant accumulation of cell mass from birth to division [3]. Additionally, recent studies on size control break down the cell-cycle in biologically relevant periods and investigate the control rules that apply during these periods to link them to molecular drivers [4]. Another emergent property of normal-sized strains is a narrow size range that is stable over generations, i.e., size homeostasis. In that regard, size convergence and homeostasis of *E.coli,* as reported by Taheri-Araghi et al., is achieved by coupling *adder* and mid-cell division processes [5].

Notably, during cell filamentation, even in clonal populations under an homogeneous environment, the variance of phenotypic traits (e.g. size, growth rate, protein expression level, etc) increases. This increase reveals a phenotypic diversity in filamentous cells. Such diversity is thought to be due to the noisy cellular biochemical reactions and can be argued to serve as a bet-hedging strategy [6].

II. Summary of Results

a). Our stochastic modeling describes quantitatively how growth and division coordinate to achieve homeostasis. Also, it sheds light on the rescue-from-filamentation process and the effect of the division positioning.

b). We reveal the conditions to satisfy *adder*-like correlations given the evident cell size compartmentalization. Also, we provide approaches to characterize the cell unit size and perform an experimental validation of our model.

d). We establish a stochastic switching model for phenotypic traits, analyze quantitatively lineage trees in terms of the number of FtzZ-rings and the switching types, and provide predictions about the fraction of different cell populations.

III. Conclusion

Our research provides a framework to understand how growth and division coordinate to achieve homeostasis and clarifies the effect of division positioning. Importantly, it shows that some *sizer* capabilities of cells associated with the divisome and chromosome duplication/segregation dynamics are compatible with the observed *adder* behavior in rod-shaped bacteria at the population level. Besides, our preliminary study of phenotypic diversity of clonal filamentous populations demonstrate trends in phenotypic switching. Overall, our work helps to understand the underlying principles of size homeostasis and phenotypic diversity during bacterial anomalous growth.

References

[1] Wehrens, M., et al. (2018). Size laws and division ring dynamics in filamentous escherichia coli cells. *Current Biology*, ***28*(6)**, 972-979.

[2] Chen, Y., Baños, R., & Buceta, J. (2018). A Markovian Approach towards Bacterial Size Control and Homeostasis in Anomalous Growth Processes. *Scientific reports*, ***8*(1)**, 9612.

[3] Amir, A. (2014). Cell size regulation in bacteria. *Physical Review Letters*, ***112*(20)**, 208102.

[4] Vargas-Garcia, C. A., Ghusinga, K. R., & Singh, A. (2018). Cell size control and gene expression homeostasis in single-cells. *Current opinion in systems biology*, ***8***, 109-116.

[5] Taheri-Araghi, S., et al. (2015). Cell-size control and homeostasis in bacteria. *Current Biology*, ***25*(3)**, 385-391.

[6] Veening, J. W., Smits, W. K., & Kuipers, O. P. (2008). Bistability, epigenetics, and bet-hedging in bacteria. *Annu. Rev. Microbiol.*, ***62***, 193-2

[1]Bioengineering Dept., Lehigh U.. E-mail: yac315@lehigh.edu
[2]Chem. and Biomol. Eng. Dept., Lehigh U.. E-mail: jab614@lehigh.edu

Using Bayesian detection theory to derive a decoder of concentration modulated signals

Chun Tung Chou[1]

***Abstract*— Cells use many different methods to encode information. In intra-cellular signalling, cells use concentration modulation (CM) to encode information in the concentration of a transcription factor. The principles behind how cells decode CM signals are not yet known. This abstract proposes a theory based on Bayesian detection theory and approximate molecular computation to derive a chemical reaction kinetics model for CM decoding. We will provide some evidence that this proposed theory holds by using the gene promotor data from Hansen and O'Shea (2013). The proposed theory can also be used to explain how gene promotor can response to a specific CM signal.**

***Index Terms*— Concentration modulation; molecular decoder; Bayesian detection; molecular computation; specificity.**

I. Introduction

In intra-cellular signalling, cells use concentration modulation (CM) to encode information in the concentration of a transcription factor (TF). This abstract presents a theory based on Bayesian detection theory, time-scale separation and approximate molecular computation to derive a chemical reaction kinetics model for CM decoding. We also present some supporting evidence for this theory.

II. Results

A. Bayesian detection for CM

In [1], we used Bayesian detection theory and renewal theorem to derive a CM deocder.

Problem statement: Let S be a TF which can turn an inactive gene promotor X into its active state X_*. We model the interactions between S and X using:

$$\mathrm{S} + \mathrm{X} \xrightarrow{g_+} \mathrm{S} + \mathrm{X}_* \tag{1a}$$

$$\mathrm{X}_* \xrightarrow{g_-} \mathrm{X}, \tag{1b}$$

where g_+ and g_- are reaction rate constants. We model these reactions by using chemical master equation. Let $x_*(t)$ be the number of X_* molecules at time t and $\mathcal{X}_*(t)$ be the continuous-time history of $x_*(\tau)$ for $\tau \in [0, t]$.

Let $u(t)$ denote the concentration of S at time t. We model $u(t)$ as a deterministic CM signal with an unknown concentration parameter a and we consider the problem of using the history $\mathcal{X}_*(t)$ to infer this unknown concentration a. For analytical tractability, we assume $u(t)$ is an ON-OFF rectangular pulse of duration d whose temporal profile is: $u(t) = a$ for $t \leq d$ and $u(t) = a_0$ for $t > d$ where a_0 is the basal concentration of the TF. We view the decoding of CM as an inference problem that computes the posteriori probability $\mathbf{P}[a|\mathcal{X}_*(t)]$.

[1]School of Computer Science and Engineering, UNSW Sydney, NSW 2052, Australia. Email: c.t.chou@unsw.edu.au

Solution: By using Bayesian theory, time-scale separation and renewal theorem, we show in [1] that the positive log-posteriori probability $\tilde{L}_a(t) = [\log(\mathbf{P}[a|\mathcal{X}_*(t)])]_+$ (where $[w]_+ = \max(w, 0)$) can be approximately computed by:

$$\frac{d\tilde{L}_a(t)}{dt} = g_-\, x_*(t) \times \left\{ \left[\log(a) - \frac{a}{u(t)} \right]_+ \right\}. \tag{2}$$

B. Applying to gene promotor data

We hypothesize that some gene promotors indirectly implement (2) where each gene promotor "calculates" $\tilde{L}_a(t)$ for a particular value of a. For a given a, we can use (2) to derive a chemical reaction kinetics model. We do that by approximating the factor in $\{\}$ in (2) by a Hill function, i.e.

$$\left\{ \left[\log(a) - \frac{a}{q} \right]_+ \right\} \approx \frac{hq^n}{H^n + q^n} \text{ for } q > \frac{a}{\log(a)} \tag{3}$$

where h, H and n are Hill function parameters that depend implicitly on the value of a. This approximation allows us to identify the right-hand side of (2) as the reaction rate of a transcription initiation process. By combining these initiation process kinetics with those of (1), mRNA and translation, we arrive at a gene expression model for CM. The essence of this model is that a CM decoder uses approximate molecular computation to calculate positive log-posteriori probability.

We fitted the derived model to the *S. cerevisiae* DCS2 gene promotor data from [2]. These data were obtained from exciting DCS2 with different TF temporal dynamics, including CM. In [1], we show that our proposed model gives a better fit than the model in [2] and is more parsimonious. This provides some supporting evidence for the proposed theory. Some additional evidence is in [1].

III. Conclusions

We proposed to use Bayesian detection theory and approximate molecular computation to explain how cells decode CM signals. The theory can be used to explain the specificity of gene promotor response to a particular CM signal. The theory can also be used for other cellular signal processing scenarios: we used it in [3] to understand the persistence detection property of the coherent feedforward loop motif.

References

[1] C. T. Chou, "Designing molecular circuits for approximate maximum a posteriori demodulation of concentration modulated signals," arXiv, Tech. Rep. arXiv:1808.01543, 2018.

[2] A. S. Hansen and E. K. O'Shea, "Promoter decoding of transcription factor dynamics involves a trade-off between noise and control of gene expression," *Molecular systems biology*, vol. 9, pp. 1–14, Nov. 2013.

[3] C. T. Chou, "Detection of persistent signals and its relation to coherent feed-forward loops," *Royal Society Open Science*, vol. 5, no. 11, Nov. 2018.

The Frailty of Nonlinear Dimensionality Reduction in single cell RNA-Seq Data

Shamus M. Cooley[1], Eric J. Deeds[2], and J. Christian J. Ray[3]

***Short Abstract* — A major component in the analysis of single cell RNA sequencing (scRNA-Seq) data is Nonlinear Dimensionality Reduction (NDR). These techniques attempt to find "embeddings" in lower dimensions that preserve some notion of the underlying structure of the data. We find that existing nonlinear dimensionality techniques do not find embeddings in the topological sense, but rather generate lower dimensional representations that disrupt the local structure of the data. Our findings suggest that new NDR techniques are needed, and that existing scRNA-Seq analyses may be producing biased results in that they assume that NDR preserves local neighborhoods, with possible consequences for interpretation of cell types and developmental lineage inference.**

***Keywords* — scRNA-Seq, Nonlinear Dimensionality Reduction, t-SNE, Isomap, PCA**

I. Background

Nonlinear Dimensionality Reduction (NDR) is employed in a wide variety of fields for visualization and analysis[1,2,3]. These techniques are often collectively referred to as "manifold learning" algorithms, but little work has been done to quantitatively evaluate the distortion of the embeddings being learned. One specific case where NDR has been utilized extensively for data analysis is in single-cell RNA-Seq (scRNA-Seq)[4]. Typically, each cell is modeled as a point in a vector space, where each dimension corresponds to the expression level of a gene in the cell's transcriptome. The high-dimensional data are often embedded in two dimensions using NDR techniques such as t-SNE[5] or UMAP[6] before further clustering and analysis. Here, we attempt to quantify, in an unbiased way, the effectiveness of NDR in finding a true embedding for any given manifold.

II. Results

We began by sampling from the surface of hyper-spheres of varying dimension in higher-dimensional vector spaces. Our results showed that, for suitably low-dimensional hyperspheres, some algorithms were able to find a close approximation to an embedding of the manifold, while other algorithms were unable to solve even this simple problem.

Next, we tried our approach on limit cycles, and found that, when the manifold in question is 2-dimensional (as limit cycles are), the algorithms tested were able to find an embedding with minimal disruption of local neighborhoods.

We then tried the same approach on hyperspheres of higher dimension, and we found that none of the commonly employed techniques could reliably generate an embedding without significant disruption of local neighborhoods.

When we tried our approach on published scRNA-Seq datasets[7,8], we obtained two important findings. First, we found from the behavior of the algorithms in various embedding dimensions, that the dimensionality of the manifold in these datasets is likely in the hundreds and cannot be reliably embedded given any of the existing NDR techniques. Second, we found that these techniques disrupt 70%-90% of local neighborhoods when embedded in any dimension, and the disruption was worst in low dimensions. This finding shows that for downstream analysis of embedded data to be truly useful, a new method of dimensionality reduction must be used in order to minimize or eliminate the disruption of neighborhoods in the data.

III. Conclusion

Our findings indicate that the information lost in existing NDR algorithms is more significant than previously assumed, and that the local structure of the data, an attribute that is often crucial to downstream analysis, is often lost completely in dimensionality reduction. This finding underscores the need to re-think the current methods of scRNA-Seq data analysis in particular, and high-dimensional data analysis in general.

References

[1] Tenenbaum J, de Silva V, Langford J. (2000) A Global Geometric Framework for nonlinear dimensionality reduction. *Science* **290**, 2319-2313

[2] Roman-Rangel E, Marchand-Maillet S. (2019) Inductive t-SNE via deep learning to visualize multi-label images. *Eng Appl Artif Intell* **81**, 336-345

[3] Ding M, Fan G. (2015) Multilayer joint gait-pose manifolds for human gait motion modeling. *IEEE Trans Cybern* **45**, 2413 – 2424

[4] Trapnell C, et. al. (2014) The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32,** 381-386

[5] Van der Maaten L, Hinton GE. (2008) Visualizing high-dimensional data using t-SNE. *J Mach Learn Res* **9**, 2579-2605

[6] McInnes L, Healy J, Melville J. (2018) UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv*:1802.03426

[7] Siebert S, Farrell J, Cazet J, et. al. (2018) Stem cell differentiation trajectories in Hydra resolved at single-cell resolution. *bioRxiv*:460154

[8] Cao J, et. al. (2019) The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496-502

[1]Bioinformatics Interdepartmental PhD Program, UCLA, Los Angeles, California, USA. E-mail: shamus@ucla.edu

[2]Institute for Quantitative and Computational Biology, UCLA, Los Angeles, California, USA. E-mail: deeds@ucla.edu

[3]Computational Biology and Molecular Biosciences, University of Kansas, Lawrence, Kansas, USA. E-mail: jjray@ku.edu

Quorum sensing in mouse embryonic stem cells controls survival and growth during differentiation

Hirad Daneshpour[1,2,3], Pim van den Bersselaar[1,2], Lars-Eric Fielmich[1,2], and Hyun Youk[1,2]

***Short Abstract* — To what extent do cells need each other to reach a decision? Here, we report a discovery of quorum sensing by differentiating mouse embryonic stem cells on a macroscopic spatial scale that determines whether the cell population expands, goes extinct, or enters a stationary phase. We link these population-level growth phases to the speed and efficiency of differentiation, using neuro-ectodermal differentiation as a case study. Time-lapse microscopy revealed that quorum sensing controls the cells' growth rate but not their death rate. Our work uncovers the importance of macroscopic range of communication on the fate of individual mouse embryonic stem cells.**

***Keywords* — Quorum sensing, survival, growth, expansion, extinction, stationary phase, differentiation, mouse embryonic stem cells.**

I. INTRODUCTION

AN important question in biology is how cells communicate amongst them to coordinate their behaviors. An important challenge is determining the length-scale at which cells communicate. Unlike for mixing liquid cultures of microbes, determining whether quorum sensing can occur for relatively immotile cells and, if so, the length-scale at which it occurs are challenging. Recent studies show that some immotile mammalian cells or organs (e.g., hair follicles) can quorum sense if they are distributed over a length scale that is a few times (about 5-10 times) their sizes [1]. An open question is if quorum sensing can occur over a macroscopic length scale that involves the entire cell population.

Here, we report a discovery of quorum sensing by differentiating mouse Embryonic Stem (ES) cells over a length scale of tens of centimeters that involves the entire cell population. ES cells are ex-vivo models, cultured on cell-culture dishes, to study pluripotency (ability to become any cell type in the adult body) and differentiation (specification of a cell type). Using neuro-ectodermal differentiation as a case study, we studied how quorum sensing governs the dynamics of survival, growth, death and differentiation in mouse ES cells.

Acknowledgements: This work was funded by Netherlands Organisation of Scientific Research (NWO), European Research Council (ERC), Canadian Institute For Scientific Research (CIFAR), and European Molecular Biology Organization (EMBO).

[1]Department of Bionanoscience
[2]Kavli Institute of Nanoscience, Delft University of Technology, The Netherlands
[3]Correspondence: E-mail: H.DaneshpourAryadi@tudelft.nl

II. RESULTS

We found that the population density (number of cells / cm^2), at the time that ES cells begin to differentiate, determines whether the entire population expands, goes extinct, or remains stationary. On the contrary, we found that a population of pluripotent ES cells of any density always expands towards the carrying capacity. Specifically, when the population of differentiating ES cells has a sufficiently high density, it expands to the carrying capacity on the culture dish whereas if it has a sufficiently low density, it will go extinct (i.e., everyone will die). Intriguingly, at an intermediate density, a differentiating population enters a stationary phase for approximately a week or longer, after which the population either expands or goes extinct. Speed and efficiency of differentiation are significantly reduced in stationary phase populations compared to the expanding populations. The range of population densities that lead to the stationary phase is narrow – just a two-fold change in the initial population density can make the difference between population expanding or going extinct.

From time-lapse imaging of differentiating cell colonies, we observed that the initial population density controls the cells' growth rate but not death rate, such that the two rates nearly match when the population has an intermediate density. Indeed, we found that cells in any population density express similar levels of the pro-apoptotic factor Bax1 while cells express higher levels of the anti-apoptotic factor Bcl2 if they are from populations of higher densities. Moreover, we used time-lapse imaging to determine that the differentiating ES cells communicate over a macroscopic scale (centimeter scales).

Finally, we could rescue populations that were destined to go extinct by transplanting them into a growth medium from a high-density, expanding population. By examining various soluble factors that cells release to the environment during differentiation, we identified growth-supporting auto-inducers (family of fibroblast growth factors) as the likely mediators of quorum sensing.

III. CONCLUSION

Our work revealed a previously unknown quorum sensing in differentiating mouse embryonic stem cells that controls the growth and death of cells, and in turn influences the speed and efficiency of differentiation.

REFERENCE

[1] Chen C-C, et al. (2015) Organ-level quorum sensing directs regeneration in hair stem cell populations. *Cell* **161**, 277-290.

Theoretical Insights into Mechanisms of Channel-Facilitated Molecular Transport in the Presence of Stochastic Gating

Aram Davtyan[1] and Anatoly B. Kolomeisky[1,2,3]

***Short Abstract* — Molecular motion through pores plays a crucial role in various natural and industrial processes. We developed a discrete-state stochastic framework to analyze the molecular mechanisms of transport processes with stochastic gating. Two scenarios are specifically investigated: 1) symmetry preserving stochastic gating with free-energy changes, and 2) stochastic gating with symmetry changes but without modifications in the overall particle-pore interactions. It is found that stochastic gating can both accelerate and slow down the molecular translocation depending on the specific parameters of the system. Our theoretical analysis clarifies physical-chemical aspects of the molecular mechanisms of transport with stochastic gating.**

***Keywords* — Molecular transport, stochastic gating, channels, molecular pores.**

I. Introduction

THE molecular transport via channels is critically important in multiple biological processes where metabolites and nutrients must be moved between different cellular compartments and delivered to specific locations.[1] It is also crucial in many industrial processes, e.g., in those that involve the separation of chemical mixtures and water purification.[2,3] The importance of translocation through pores stimulated extensive theoretical studies to uncover the underlying molecular mechanisms. But many questions remain open. Specifically, most of existing theoretical studies of channel-facilitated molecular transport concentrate on investigating systems, where interactions between particles and the channel are constant over the time. However, biological cells are very dynamic non-equilibrium systems, where inter-molecular interactions frequently change as a result of passive or active regulation processes. For instance, ion channels are largely regulated in biological cells by varying the membrane potentials and by changing the dynamics of ligands binding to membrane receptors.[4] As a result, the channel can undergo significant conformational changes that might close or restrict the passage of particles through it for some periods of time. This is known as a *stochastic gating* phenomenon, and it is widely observed in biological systems.

In this work we developed a simple theory of stochastic gating for particles traveling through molecular channels using the discrete-state chemical-kinetic approach. Our goal is to understand the general features of the stochastic gating and how it can optimize the molecular transport. We specifically consider two limiting situations: 1) when the stochastic gating is associated with fluctuations in the free-energy for a pore system that is always symmetric; and 2) when the stochastic gating changes the symmetry of the interaction potential without overall modifications in the interaction strength between the molecule and the pore.

II. Conclusion

Our theoretical analysis explicitly evaluates the particle currents through the pores in terms of transition rates between various chemical states and conformation. It allows us to specifically investigate two different models of stochastic gating. In the first model, the stochastic gating leads to the changes in the translocation free energy profile but without symmetry variations. It is found that increasing the frequency of conformational transitions and the concentration gradients between different parts of the channel will always increase the particle current through the system. At the same time, varying the interaction energy between the molecules and the pores generally leads to non-monotonic behavior. A more complex dynamic behavior is observed in the second model of stochastic gating that involves symmetry variations in the free-energy translocation profile without changing the overall interactions. While increasing the concentration gradient will always accelerate the molecular fluxes, the dependence on the frequency of conformational fluctuations and on interaction energies is non-monotonic. We presented microscopic arguments to explain these observations. Importantly, in both models we do not observe phenomena similar to a resonance activation when there is an optimal rate of conformational transitions that leads to a maximal particle current.

References

[1] Lodish H, Berk A, Darnell JE, Kaiser CA, Krieger M, Scott MP, Bretscher A, Ploegh H, Matsudaira P, et al. (2008) In *Molecular cell biology* (Macmillan).

[2] Meloan CE (1999) In *Chemical separations: principles, techniques and experiments* (Wiley-Interscience).

[3] Sparreboom W, van den Berg A, and Eijkel JC (2009) "Principles and applications of nanofluidic transport" *Nature nanotechnology* **4**, 713.

[4] Hille B. et al. (2001) In *Ion channels of excitable membranes*, Vol. 507 (Sinauer Sunderland, MA).

The work was supported by the Welch Foundation (C-1559), by the NSF (CHE-1664218), and by the Center for Theoretical Biological Physics sponsored by the NSF (PHY-1427654).

[1]Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, USA. E-mail: adavtyan@rice.edu

[2]Department of Chemistry, Rice University, Houston, Texas 77005-1892, USA. E-mail: tolya@rice.edu

[3]Department of Chemical and Biomolecular Engineering, Rice University, Houston, Texas 77005-1892, USA

Tuning Spatial Profiles of Selection Pressure to Modulate the Evolution of Antibiotic Resistance

Maxwell G. De Jong[1], Kevin B.Wood [1,2]

***Abstract*— To better understand the impact of spatial heterogeneity on the evolution of antibiotic resistance, we develop a toy model of stochastic microbial dynamics and use a mean first passage time calculation to investigate how spatial profiles of selection pressure impact the time to fixation of a resistant allele. We observe that spatial heterogeneity can accelerate or decelerate resistance with an arbitrary initial mutant subpopulation, depending on the mutation and migration rates of the microbe, and incorporating a fitness cost amplifies this effect. Finally, we investigate analytical approximations that allow us to explore larger systems.**

***Index Terms*— Spatial heterogeneity, microbial evolution, non-equilibrium statistical mechanics, antibiotic resistance.**

I. Background

EMERGENT behavior in communities complicates the evolution of antibiotic resistance. Several experiments have demonstrated that spatial drug gradients can significantly speed the evolution of resistance in simple microbial communities [1], [2]. While theoretical work has been performed to understand these specific results [3], there is still relatively little known about resistance in complex spatial profiles of selection pressure, which can be present in biological systems due to any number of spatially-varying quantities.

In this work, we develop a toy model of microbial evolution in a spatially-extended system composed of three connected microhabitats. Cells can exist as wild-type cells or drug-resistant mutants in each of these microhabitats, and cells migrate between neighboring microhabitats at a fixed migration rate. We enumerate over the states of the system to write down a master equation, and we solve for the mean fixation time using a mean first passage time commonly used in statistical physics. The mean fixation times are affected by the mutation rate, migration rate, and spatial distribution of selection pressure. Fixation times for different selection pressure landscapes are compared to those obtained with a spatially-homogeneous landscape. Importantly, the mean selection pressure is conserved across different landscapes. Because this method allows average fixation times to be calculated from arbitrary initial states of our system, the role of spatial heterogeneity in fixation can readily be calculated for any initial mutant subpopulation.

[1]Department of Physics, University of Michigan. Email: mgdejong@umich.edu

[2]Department of Biophysics, University of Michigan. Email: kbwood@umich.edu

II. Results

We have found that the fixation time for an initially wild-type population can be significantly accelerated or decelerated by varying the spatial distribution of selection pressure, even with the spatially-averaged selection pressure fixed [4]. When the migration rate is sufficiently large relative to the mutation rate, the fixation time is accelerated by spatial heterogeneity and can be approximated as the minimum of three exponentially-distributed random variables. When the migration rate is small, heterogeneity slows fixation times, which can be approximated as the maximum of three exponentially-distributed random variables. We find similar results for arbitrary initial mutant subpopulations, and we look at how the optimal selection distribution to maximally slow fixation depends on this initial subpopulation.

We expand upon this work by assuming that the mutation conferring resistance comes at some fitness cost. This leads to qualitatively different results—because genetic drift can now be biased against the mutants in some microhabitats, fixation is not guaranteed to be reached in the deterministic limit. We observe that the impact of spatial heterogeneity on fixation times is greatly amplified, but much of our intuitive understanding from the system without a fitness cost still holds. Additionally, we look at approximating our fixation times outside of either of the limiting cases for the migration and mutation rates. Without any approximation, the number of equations scales as a power law in the number of cells and exponentially in the number of microhabitats.

III. Conclusion

We demonstrate spatial distribution of selection pressure can significantly impact the times to fixation in a toy model even while the spatially-averaged selection is fixed. This work helps to better understand the evolution of antibiotic resistance in spatially-structured populations, and our results may lay the groundwork for optimized, spatially-resolved drug dosing strategies for mitigating antibiotic resistance.

References

[1] Q. Zhang, G. Lambert *et al.*, "Acceleration of emergence of bacterial antibiotic resistance in connected microenvironments," *Science*, vol. 333, pp. 1764–1767, 2011.

[2] M. Baym, T. D. Lieberman *et al.*, "Spatiotemporal microbial evolution on antibiotic landscapes," *Science*, vol. 353, pp. 1147–1151, 2016.

[3] P. Greulich, B. Waclaw, and R. J. Allen, "Mutational pathway determines whether drug gradients accelerate evolution of drug-resistant cells," *Phys. Rev. Lett.*, vol. 109, p. 088101, Aug 2012.

[4] M. G. De Jong and K. B. Wood, "Tuning spatial profiles of selection pressure to modulate the evolution of drug resistance," *Phys. Rev. Lett.*, vol. 120, p. 238102, Jun 2018.

Constitutive splicing and economies of scale in gene expression

Fangyuan Ding[1,2], Michael B. Elowitz[1,2*]

[1]Division of Biology and Biological Engineering, Caltech, 1200 E. California Blvd. Pasadena, CA 91125, USA
[2]Howard Hughes Medical Institute
*melowitz@caltech.edu

Abstract

In eukaryotic cells, many introns are constitutively, rather than alternatively, spliced, and therefore do not contribute to isoform diversification. It has remained unclear what functional roles such constitutive splicing could provide. To explore this issue, we asked how splicing affects the efficiency with which individual pre-mRNA transcripts are productively processed across different gene expression levels. We developed a quantitative single-molecule FISH-based method to quantify splicing efficiency at the transcription active site in single cells. Using this method, we found both natural and synthetic genes exhibit an 'economy of scale' behavior in which splicing efficiency increased with transcription rate, rather than decreasing as expected for a standard enzymatic process. Correlations between splicing efficiency and spatial proximity to nuclear speckles could explain this counterintuitive behavior. Functionally, economy of scale splicing represents a non-linear filter that amplifies the expression of genes when they are more strongly transcribed. These results indicate that constitutive splicing plays an active functional role in modulating gene expression.

Tensor train solution to the equilibrium of the Chemical Master Equation

Trang N. Dinh[1] and Roger B. Sidje[1]

***Abstract*— Solving the stationary solution of the Chemical Master Equation is difficult due to the curse of dimensionality. By converting the problem in matrix format to quantized tensor train (QTT) format, the solution is presented in a compressed form that scales linearly with the dimension. We investigate tensor-based approaches to approximate the probability distribution in a long run, and show its efficiency as more chemical species are introduced to the system.**

***Index Terms*— chemical master equation, stationary solution, quantized tensor train format, tensor-based approaches**

I. Background

The Chemical Master Equation (CME) governs the changes in a biological pathway with multiple species as a system of linear equations that is frequently represented in matrix form:

$$\dot{\mathbf{p}}(t) = \mathbf{A}\mathbf{p}(t), \;\; t \in [0, t_f]$$

where $\mathbf{A}$ is the $n \times n$ transition rate matrix of the Markov chain underlying the CME, and $\mathbf{p}(t)$ the $n \times 1$ probability vector with each component p_i the probability of finding the system in state i at time t. The probability vector at t_f is given by:

$$\mathbf{p}(t_f) = \exp(t_f \mathbf{A})\mathbf{p}(0).$$

The derivative, $\dot{\mathbf{p}}$, will simply be the zero vector when the system is in statistical equilibrium. Hence, we can find the stationary probability vector by solving the linear system

$$\mathbf{A}\overset{\infty}{\mathbf{p}} = \mathbf{0}.$$

The size of the CME increases exponentially with the number of species. Therefore, there is a need to store the CME in a more compact fashion for realistic biological models. The tensor format is one such approach.

II. Tensor-based Approximation of The Stationary Solution to the CME

Any $I_1 \times I_2 \times \cdots I_N$ tensor $\mathcal{X}$ can be represented in tensor train format (TT-format) [1]

$$\mathcal{X}(i_1, \cdots, i_N) = \mathcal{G}_1(i_1) \cdots \mathcal{G}_N(i_N),$$

where $\mathcal{G}_k(i_k) \in R^{r_{k-1} \times r_k}, i_k \in 1, \cdots, I_k$. With the assumption that $I_s = 2^{L_s}$, $\mathcal{X}$ can be reshaped into a tensor $\mathcal{Y}$ of order $L := L_1 + \cdots + L_N$ and size $2 \times \cdots \times 2$. The quantized tensor train format of $\mathcal{X}$ (QTT format of $\mathcal{X}$) is then defined as the TT format of the reshaped tensor $\mathcal{Y}$. Let $\mathbf{p}$ be the current transient solution associated with the hyper-rectangle $\mathbf{H} = \{0, \cdots, 2^{L_1} - 1\} \times \cdots \times \{0, \cdots, 2^{L_N} - 1\}$. Then $\mathbf{p}$ can be interpreted as a multidimensional array of dimension $I_1 \times I_2 \times \cdots I_N$ and therefore be written in the QTT format

$$\mathbf{p}(i_1, \cdots, i_N) = \mathbf{P}_1^{(1)}(i_1^{(1)}) \cdots \mathbf{P}_{L_1}^{(1)}(i_{L_1}^{(1)}) \cdots \mathbf{P}_N^{(1)}(i_N^{(1)}) \cdots \mathbf{P}_{L_N}^{(1)}(i_{L_N}^{(1)}).$$

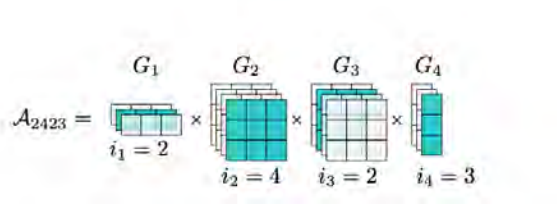


Fig. 1. An illustration of the TT-format for a tensor $\mathcal{A}$ in four dimensions.

The Adaptive FSP-QTT algorithm [2] shows substantial savings in excecution time on top of the impressive savings in memory enabled by the QTT/TT format. Tensor-based approaches have also been applied to the study of the CME's stationary solution [3]. In this work, we present methods to approximate the stationary solution in QTT format, such as inexact uniformization [4] and iterative techniques [5], to approximate the stationary solution in QTT format and compare their computational efficiency to more traditional methods.

References

[1] I. V. Oseledets, "Tensor-train decomposition," *SIAM J. Sci. Comput.*, no. 33(5), 2295-2317, 2011.
[2] H. D. Vo and R. B. Sidje, "An adaptive solution to the chemical master equation using tensors," *J. Chem. Phys.*, no. 147(4):044102, 2017.
[3] V. Kazeev and C. Schwab, "Tensor approximation of stationary distributions of chemical reaction networks," *SIAM J. Matrix Ana. Appl.*, vol. 36, no. 3, pp. 1221–1247, 2015.
[4] R. B. Sidje, K. Burrage, and S. Macnamara, "Inexact uniformization method for computing transient distributions of markov chains," *SIAM J. Sci. Comput*, vol. 29, no. 6, p. 25622580, 2007.
[5] Y. Saad, "Iterative methods for sparse linear systems," 2003.

[1]The University of Alabama, Department of Mathematics, P.O. Box 870350, Tuscaloosa, AL 35487, USA. Email: tndinh@crimson.ua.edu, roger.b.sidje@ua.edu.

Parametrizing a Quantitative Model of Genetic Circuit Function with a Single Experiment

Hamid Doosthosseini[1], Amin Espah Borujeni[1], Jing Zhang[1], and Christopher A. Voigt[1]

Developing applicable quantitative models is faced with the challenge of parametrization, especially for systems with limited prior characterization. Rapid advances in synthetic biology have led to interest in implementing new genetic circuits in previously uncharacterized host organisms. In this work, we present a methodology to fully parametrize a descriptive model of a circuit's transcription and translation profiles using a transcriptomic data set exclusively. This method would not require any a priori information about the host and, as such, could serve to characterize and potentially predict the dynamic performance of a genetic circuit in an understudied host.

I. Purpose

ENGINEERING genetic regulatory networks capable of performing complex logic operations was made possible through the development of modular genetic constructs or "gates" [1-3]. Bottom-up design in this manner is limited an requires models able to describe the performance of each gate and compiled constructs [2,3]. Biophysical kinetic models are great candidates to this end, however, there is a delicate balance in quantifying biophysical parameters without overfitting, over-parametrizing, or relying on multiple literature sources. It would be ideal to fully parametrize such a model from a single set of experiments.

Transcriptomics data, namely RNA sequencing and ribosomal profiling, have the capacity to produce genome-wide data, and can be an ideal platform to parametrize genetic parts, context effects, and translation [3,4].

In this work, we demonstrate that a biophysical model of RNAP flux across a 3-input logic genetic construct, can be fully parametrized from a single set of experiments. Furthermore, this model captures the steady-state transcription and translation profiles and, assuming dilution rate, predicts dynamics. This work also elucidates a number of genetic context effects, suggesting improvements are possible in designing the topology of a circuit.

II. Summary of Results

A system of linear ODEs were derived using biophysical and kinetic models of transcription, transcriptional regulation and translation. 5 parameters were found sufficient to fully capture the steady state behavior of the circuit, see table 1. The model reconstructs transcription and translation profiles accurately (fig. 1), as well as predicting dynamic responses.

Read-through of RNAP flux from upstream was found to significantly impact repressor expression and modeling these context effects improved model accuracy. Additionally, read-through was found to roadblock repressor binding to promoter rendering repression less effective, significantly impacting tandem promoter functionality in circuit designs.

Table 1. List of parameters describing genetic parts

Part	Parameters
Terminator	Strength
Ribozyme	Stability coefficient
Gene	Relative translation efficiency
Promoter	Maximal activity
Promoter	Binding coefficient of transcription factor

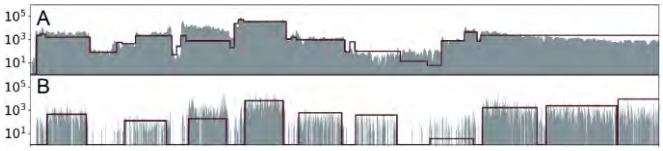


Fig 1. Original (grey) and reconstructed (brown) profiles of (A) transcription and (B) translation across the plasmid

III. Conclusion

RNA sequencing and Ribosomal profiling are sufficient for full parametrization of a descriptive model, allowing reconstruction of transcription and translation profiles as well as prediction of dynamics without any a priori information about the host organism. Genetic context is shown to significantly impact circuit performance and modeling this using transcriptomics data remarkably improves accuracy

References

[1] Anderson, J. C., et. al. "Environmental signal integration by a modular AND gate." *Molecular systems biology* 3.1 (2007): 133.

[2] Nielsen, A. A. K., et al. "Genetic circuit design automation." *Science* 352.6281 (2016): aac7341.

[3] Gorochowski, T. E., et al. "Genetic circuit characterization and debugging using RNA-seq." *Molecular systems biology* 13.11 (2017): 952.

[4] Li, G., et al. "Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources." *Cell* 157.3 (2014): 624-635.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Defense Advanced Research Projects Agency (DARPA), the Department of Defense, or the United States Government. This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) and the Air Force Research Laboratory under Contract No. [FA8750-17-C-0229]

[1]Synthetic Biology Center, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

Presenting author: Hamid Doosthosseini, Email: hdoosth@mit.edu

Dynamics, Noise, and Antibiotic Resistance in Single Cells

Mary J. Dunlop

The majority of our understanding on antimicrobial drug resistance comes from studies on the genetic changes that cause it, however bacteria can also transiently survive antibiotic exposure even without permanent genetic changes. Using a combination of time-lapse microscopy experiments and stochastic modeling I will show how *E. coli* bacteria use feedback to generate dynamics and noise in expression of a key regulatory protein, providing transient antibiotic resistance at the single-cell level. In addition, I will discuss how expression of resistance genes can predispose cells towards mutation. Our findings show that even in the absence of antibiotic exposure, cells that are transiently resistant are more mutation prone, suggesting that these transient mechanisms may act as a stepping stone towards higher levels of drug resistance. These results are significant because they reveal important dynamic information about the period over which transient resistance develops and ultimately how it can lead to permanent genetic changes encoding multidrug resistance.

Acknowledgements: This work was funded by NIH grant XX00000. This footnote is optional.
[1]Department of Systems Biology, University of Quantitative Biology, Other address information. E-mail: author@place.edu
[2]Mathematical Biology Center, BioPark USA. E-mail: two@place.gov
[3]Another Dept, Another Institution, Address, E-mail: three@place.com

Graph Traversal Edit Distance

Ali Ebrahimpour-Boroojeny[1], Akash Shrestha[1], Ali Sharifi-Zarchi[1], Suzanne R. Gallagher[1], S. Cenk Sahinalp[2], and Hamidreza Chitsaz[1]

***Abstract*— Many problems in applied machine learning deal with graphs (also called networks), including social networks, security, web data mining, protein function prediction, and genome informatics. In this project, we introduce a new measure for comparing two graphs which we call graph traversal edit distance (GTED), and we give the first polynomial time algorithm for it. GTED is motivated by and provides the first mathematical formalism for sequence co-assembly and *de novo* variation detection in bioinformatics. We demonstrate that GTED admits a polynomial time algorithm using a linear program in the graph product space that is guaranteed to yield an integer solution.**

***Index Terms*— Graph Kernels, Differential Genome Assembly, Sequence Co-assembly, de Bruijn Graphs, Assembly graphs, Graph Comparison**

I. Purpose

NEtworks, or graphs as they are called in mathematics, have become a common tool in modern biology. Biological information from DNA sequences to protein interaction to metabolic data to the shapes of important biological chemicals are often encoded in networks.

One goal of studying these networks is to compare them. We might want to know whether two DNA assembly graphs produce the same final sequences or how close the protein interaction networks of two related species are. Such comparisons are difficult owing to the fact that determining whether two graphs have an identical structure with different labels or vertex ordering is an NP-complete problem. Therefore, any comparison will need to focus on specific aspects of the graph.

Here, we present the notion of *graph traversal edit distance (GTED)* [1], a new method of comparing two networks. Informally, GTED gives a measure of similarity between two directed Eulerian graphs with labeled edges by looking at the smallest edit distance that can be obtained between strings from each graph via an Eulerian traversal. GTED was inspired by the problem of *differential genome assembly*, determining if two DNA assembly graphs will assemble to the same string. In the differential genome assembly problem, we have the de Bruijn graph representations (assembly graphs) of two (highly) related genome sequence data sets (e.g. one from a cancer tissue and the other from the normal tissue of the same individual). Differential genome assembly has been introduced to bioinformatics in two flavors: (i) *reference genome free* version [2], [3], [4], [5], [6], and (ii) *reference genome dependent* version, which, in its most general form, is NP-hard [7]. Both versions of the problem are attracting significant attention in biomedical applications (e.g. [8], [9]) due to the reduced cost of genome sequencing (now approaching $1000 per genome sample) and the increasing needs of cancer genomics where tumor genome sequences may significantly differ from the normal genome sequence from the same individual through single symbol edits (insertions, deletions, and substitutions) as well as block edits (duplications, deletions, translocations, and reversals).

In addition to comparing assembly graphs, GTED can be used as a graph kernel for a number of classification problems that involve graphs. GTED is the first mathematical formalism in which global traversals play a direct role in the graph metric. In this paper, we give a polynomial time algorithm using linear programming that is guaranteed to yield an integer solution.

II. Experiments and Results

We use GTED as a graph kernel and evaluate its performance in SVM classification over four benchmark datasets. Our new kernel is able to get higher accuracy on two of the datasets, Mutag [10] and Enzymes [11], compared to other common graph kernels. We also use GTED for clustering of viral genomes obtained from *de novo* assembly of next-generation sequencing reads. Our new approach is able to successfully cluster them into their corresponding genera.

III. Conclusion

We introduced a new measure for comparing two graphs that was motivated by the problem of differential genome assembly. We also provided a polynomial algorithm for computing this measure. Although this method is not applicable to large graphs, it can form the mathematical basis for scalable heuristic comparison of full-size large genomes in the future.

References

[1] A. E. Boroojeny, A. Shrestha, A. Sharifi-Zarchi, S. R. Gallagher, S. C. Sahinalp, and H. Chitsaz, "Gted: Graph traversal edit distance," in *International Conference on Research in Computational Molecular Biology*. Springer, 2018, pp. 37–53.

[2] Y. L. *et al.*, "Structural variation in two human genomes mapped at single-nucleotide resolution by whole genome *de novo* assembly," *Nature Biotechnology*, vol. 29, pp. 723–730, 2011.

[3] N. S. Movahedi, E. Forouzmand, and H. Chitsaz, "De novo co-assembly of bacterial genomes from multiple single cells," in *IEEE Conference on Bioinformatics and Biomedicine*, 2012, pp. 561–565.

[1]Department of Computer Science, Colorado State University, Fort Collins. Email: chitsazlab@lists.colostate.edu

[2]Department of Computer Science, Indiana University, Bloomington. Email: cenksahi@indiana.edu

[4] Z. Iqbal, M. Caccamo, I. Turner, P. Flicek, and G. McVean, "De novo assembly and genotyping of variants using colored de bruijn graphs," *Nat Genetics*, vol. 44, pp. 226–232, 2012.

[5] Z. Taghavi, N. S. Movahedi, S. Draghici, and H. Chitsaz, "Distilled single-cell genome sequencing and de novo assembly for sparse microbial communities," *Bioinformatics*, vol. 29, no. 19, pp. 2395–2401, Oct 2013.

[6] N. S. Movahedi, M. Embree, H. Nagarajan, K. Zengler, and H. Chitsaz, "Efficient synergistic single-cell genome assembly," *Frontiers in Bioengineering and Biotechnology*, vol. 4, no. 42, 2016.

[7] F. Hormozdiari, I. Hajirasouliha, A. McPherson, E. Eichler, and S. C. Sahinalp, "Simultaneous structural variation discovery among multiple paired-end sequenced genomes," *Genome Research*, vol. 21, pp. 2203–2212, 2011.

[8] C. Mak, "Multigenome analysis of variation (research highlights)," *Nature Biotechnology*, vol. 29, 2011.

[9] S. Jones, "True colors of genome variation (research highlights)," *Nature Biotechnology*, vol. 30, 2012.

[10] A. K. Debnath, R. L. Lopez de Compadre, G. Debnath, A. J. Shusterman, and C. Hansch, "Structure-activity relationship of mutagenic aromatic and heteroaromatic nitro compounds. correlation with molecular orbital energies and hydrophobicity," *Journal of medicinal chemistry*, vol. 34, no. 2, pp. 786–797, 1991.

[11] K. M. Borgwardt, C. S. Ong, S. Schönauer, S. V. N. Vishwanathan, A. J. Smola, and H.-P. Kriegel, "Protein function prediction via graph kernels," *Bioinformatics*, vol. 21, no. 1, pp. 47–56, Jan. 2005. [Online]. Available: http://dx.doi.org/10.1093/bioinformatics/bti1007

Gene expression differences explain heterogeneity in calcium signaling response to ATP

Robert Foreman and Roy Wollman

***Short Abstract* — Single-cells exhibit marked heterogeneity in calcium signaling after stimulation with ligand. However, it is unknown how much of the variability observed comes from differences in gene expression, or differences at the level of the signaling network such as ligand receptor activity differences. Here we measure the live-cell calcium response and 300+ genes in the calcium signaling network from the same cells. We find that cell state differences in the gene expression of calcium signaling genes explain some, but not all of the heterogeneity observed in the dynamic calcium response.**

***Keywords* — cell state, single-cell, MERFISH, Calcium Signaling, and qBio**

I. Introduction

Single cells transcriptional measurements are revealing putative cell types and states with increasing details. [1] However, the degree by which such putative clusters in expression space represent cell states with distinct phenotypes is unclear, especially when the separation between transcriptional states is small or depends on highly non-linear dimensionality reduction methodology.

To address this question we focused on calcium response of epithelial mammary cells to external cues, an experimental model with prior indication of the existence of distinct yet similar cell states. [2] We performed single cell joint measurement of calcium dynamics and the expression of 300+ genes from calcium signaling network using merFISH protocol. [3]

II. Results

We show there are many putative definition of cell states based solely on expression or calcium dynamics making each dataset alone insufficient to determine the existence of underlying cell states. However, by combining results from transcriptional and phenotypic analysis the underlying cell states become clear showing a distinct separation into two distinct cell states.

References

[1] Kumar, M. P. *et al.* Analysis of Single-Cell RNA-Seq Identifies Cell-Cell Communication Associated with Tumor Characteristics. *Cell Rep.* **25**, 1458–1468.e4 (2018).

[2] Yao, J., Pilko, A. & Wollman, R. Distinct cellular states determine calcium signaling response. *Mol. Syst. Biol.* **12**, 894 (2016).

[3] Moffitt, J. R. *et al.* High-performance multiplexed fluorescence in situ hybridization in culture and tissue with matrix imprinting and clearing. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 14456–14461 (2016).

[1]Bioinformatics and Systems Biology, UCSD E-mail: rforeman@ucsd.edu

[2]Department of Biochemistry, UCLA. E-mail: rwollman@ucla.edu

Intrinsic gene variability is near Poisson

Robert Foreman and Roy Wollman

Short Abstract — **Mammalian gene expression is known to be highly variable (over dispersed relative to Poisson expectation). Previous studies have indicated that a major source of this over dispersion comes from intrinsic noise due to transcriptional bursting. However, previous works did not exhaustively consider all sources of extrinsic factors in their decomposition of intrinsic and extrinsic noise sources. Here we show near Poisson intrinsic noise for most genes in a set of 150 genes related to calcium signaling in mammalian MCF10A breast epithelial cells.**

With the increasing number of single-cell studies revealing detailed organization of cells into type and sub-states within each type, understanding the relative contribution of different sources of variable helps elucidate how these different cell types are created and maintained.

Keywords — **intrinsic/extrinsic noise, MERFISH, Calcium Signaling, and qBio**

I. Introduction

Gene expression variability is fundamental to biological systems. In multicellular organisms, heterogeneity between different cells confers specialized function giving rise to complexity in whole-system behavior. Landmark works investigating this heterogeneity in bacterial cells laid a foundation for decomposing variability into intrinsic (local) and extrinsic (external) sources; where intrinsic variability derives from stochastic events local to the transcriptional site that manifest as noisy allele specific differences in expression, and extrinsic variability arises from non-local 'systematic' differences affecting both alleles or multiple genes, such as varying concentrations of transcription factors between cells. [1-2]

Classically extrinsic variability has been determined by measuring covariates of a gene's expression and determining the amount of variance explained by that covariate. Elowitz and O'Shea, in bacteria and yeast respectively, cloned the same gene in two colors were one color's expression was the covariate for the other gene. [2-3] Similar work in mammalian systems concluded that gene expression was more variable than the naïve expectation of a Poisson process. Two-state, transcriptional bursting, models of gene expression were proposed to explain this over dispersion. [4]

Real-time measurement of gene expression using the MS2 system appeared to show periods of fluctuating transcriptional rates on the time scale of minutes to hours between bursts. These measurements were corrected for some extrinsic covariates such as cell-cycle and cellular volume, but exhaustive consideration of all extrinsic factors were not considered. [5-6]

Here we measure 150 genes' expression values with a highly sensitive and accurate smMERFISH method. We then use measured covariates such as cell cycle, cell volume, MCF10A differentiation markers, and dynamical calcium response to ATP stimulation to decompose intrinsic and extrinsic variability. We also infer unknown extrinsic factors through PCA fitting of gene-to-gene covariance, and consider the scale of the residual variance arising from intrinsic sources.

II. Results

Residual intrinsic variability for most of the 150 genes is near Poisson. Transcriptional bursting and intrinsic noise play a less significant role in generating gene expression variability than previously estimated. This difference likely arises from previous methods inability to account for enough sources of extrinsic variability, or filtering/compensation of transcription bursts such that the scale of induced intrinsic variance in near Poisson.

[1]Bioinformatics and Systems Biology, UCSD E-mail: rforeman@ucsd.edu

[2]Department of Biochemistry, UCLA. E-mail: rwollman@ucla.edu

REFERENCES

[1] Dar, R. D., Razooky, B. S., Weinberger, L. S., Cox, C. D. & Simpson, M. L. The Low Noise Limit in Gene Expression. *PLoS One* **10**, e0140969 (2015).

[2] Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).

[3] Raser, J. M. & O'Shea, E. K. Control of stochasticity in eukaryotic gene expression. *Science* **304**, 1811–1814 (2004).

[4] Tantale, K. *et al.* A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat. Commun.* **7**, 12248 (2016).

[5] Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. & Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **4**, e309 (2006).

[6] Suter, D. M. *et al.* Mammalian genes are transcribed with widely different bursting kinetics. *Science* **332**, 472–474 (2011).

Modeling protein-membrane interactions through an efficient continuum model of lipids

Yiben Fu[1], and Margaret E. Johnson[1]

***Abstract* — Protein self-assembly is critical in driving membrane remodeling, such as in viral bud formation. Reaction-diffusion is a powerful technique to simulate the process at the cell-scale. An important issue is how to accurately model the binding of proteins with membranes. Here we present a model for molecules binding to membrane surfaces, with lipids distributed continuously on membrane. This method efficiently captures the influence of different populations of lipids on binding kinetics, and can account for heterogeneity in the surface properties. Importantly it can also describe 2D interactions.**

***Keywords* — Protein self-assembly, Reaction-diffusion simulations, Binding to membranes.**

I. Purpose

THE self-assembly of cytosolic proteins on membranes plays a critical role in driving membrane remodeling in vesicle trafficking and viral bud formation [1,2]. Simulating these processes at the cell-scale is challenging due to their relatively long time-scales and coupling to nonequilibrium processes. Structure-resolved reaction-diffusion is a powerful technique for performing these types of simulations [3,4]. An important question is how to accurately and efficiently model the binding and interactions of single protein molecules with and on membranes, at the cell-scale and biologically-relevant time-scales. Here, we present a model for molecules binding to membrane surfaces, with lipids distributed continuously on membrane.

II. Methods

Many proteins bind lipid bilayers with high-specificity to specific lipids, such as phosphatidyl-inositols (PI). Propagation of individual lipids in single-particle reaction-diffusion simulations, however, is computationally expensive. To alleviate this, we propose that a more efficient approach is to replace the lipids by a density distribution. This approach will reduce the spatial detail of the lipid distribution, but will eliminate the need to propagate the numerous particles on the membrane surface, which are often more numerous than the solvated molecules.

This approach replaces pairwise reaction probabilities between solution-phase molecules and membrane-bound particles with a reaction probability between the solution molecule and the field on the surface. To do so, we integrate over the patch of surface that is accessible to the solution-phase molecule, and the reaction probability across this patch is scaled by the mean density of lipid particles.

III. Results

The development and comparison with theory and alternative models of binding shows that this method is highly efficient.

A. Comparison to different approaches

This new method produces the same kinetics as the single-particle approach [3], and captures the influence of a small population of lipids for enabling membrane binding.

Compared to related models of surface adsorption [5], our model decouples the binding affinity and concentration, which provides more flexibility for controlling surface binding properties. We further show that our method can account for heterogeneity in the surface properties, which can occur, for example, due to lipid phase separation.

B. Localization of proteins to 2D

Importantly, this method also allows account for 2D problems that protein molecules which are localized on the surface, can bind to each other through a protein-protein interaction. This effective 2D interaction is important for increasing stability of multi-protein complexes on the surface that contain multiple lipid binding sites.

C. Further application

This method will be demonstrated for curved membrane surfaces. Thus, we can present the progress on implementing a dynamic membrane that is integrated with an accurate propagator for reaction-diffusion dynamics in solution, on surfaces, and transitioning between them.

IV. Conclusion

We develop a method for molecules binding to membrane surfaces by describing the lipids as a continuous distribution on surfaces. This method is efficient and can depict both the surface properties and the kinetic process of binding to membrane surfaces.

References

[1] Avinoam O, et al (2015) *Science* **348**, 1369-1372.
[2] Votteler J, Sundquist WI (2013) *Cell Host Microbe* **14**, 232-241.
[3] Johnson ME, Hummer G (2014) *Phys Rev X* **4**, 031037.
[4] Johnson ME (2018) *J Phys Chem B* **122**, 11771-11783.
[5] Andrews SS (2009) *Phys Biol* **6**, 046015.

Acknowledgements: This work was funded by NSF grant 1753174.
[1]Department of Biophysics, The Johns Hopkins University, Baltimore, MD, USA. E-mail: margaret.johnson@jhu.edu

Deciphering Whole-Organism Perturbations with Multiplexed Single-Cell RNA-Seq

Jase Gehring[1], Brady Weissbourd[1], David Anderson[1] and Lior Pachter[1,2]

***Short Abstract* — Multiplexed single-cell RNA-seq is emerging as a powerful platform for parallel interrogation of complex perturbation experiments. Using the medusa (jellyfish) stage of the Cnidarian *Clytia hemisphaerica*, we have recently completed the first whole-organism multiplexed scRNA-seq experiments. We compared five healthy animals to five starved specimens and obtained the first cell atlas for a medusa-stage Cnidarian. In addition to rich cell type information and developmental trajectories, we also quantified gene expression changes in response to starvation for every cell type in the animal. Organism-wide perturbation experiments like ours promise to revolutionize our understanding of coordinated cellular dynamics across entire animals.**

***Keywords* — Systems Biology, Single-Cell Biology, Non-Model Organisms, Bioinformatics, Biostatistics, Pertubations**

I. Background

TECHNOLOGICAL advances in parallel cell barcoding have driven impressive scaling in single-cell RNA-sequencing (scRNA-seq) experiments, with commercial and academic platforms now capable of processing 10^4-10^6 cells in a single experiment [1-3]. At this scale, it is possible to combine many barcoded samples into a single batch, and accordingly many methods for sample multiplexing have been introduced in the past year [4-6]. We have developed a universal sample multiplexing method based on chemical deposition of activated oligonucleotides (ClickTags) on fixed cell suspensions [7]. In a proof-of-concept experiment, we subjected neural stem cells to combinations of growth factors at varying concentrations and analyzed the resulting heterogeneous populations on the 10x Chromium system. We have now extended this approach to label cell suspensions derived from entire animals.

II. Results

Individual *Clytia hemisphaerica* medusae (jellyfish) were dissociated, labeled with ClickTags, and analyzed by multiplexed scRNA-seq, comparing five starved animals with five fed animals. In a single experiment, we were able to produce a cell atlas for this non-model organism, identify the origin and developmental trajectories of a number of unique cell types, characterize gene expression variance across individuals, and perform cell-type specific differential expression for the fed vs. starved conditions. This experiment illustrates the wealth of information afforded by multiplexed whole-organism scRNA-seq, including a systems-level view of differential gene expression in response to experimental stimuli at cell-type resolution.

Analyzing such a dataset requires a suite of bioinformatics tools and novel statistical approaches. We developed an ultra-fast demultiplexing algorithm, dubbed kITE, to assign cells to their animal-of-origin. RNA velocity [8] and pseudotime [9] analyses were used to infer developmental trajectories for many cell types, and a novel test for cluster-specific marker genes was employed that takes advantages of the biological replicates present in the experiment. We identified 36 distinguishable cell types including a complex developmental program in which neural precursors differentiate into either stinging cells or one of a range of neurons that make up a complex Cnidarian nervous system. Finally, the starvation response was characterized and validated by fluorescence *in situ* hybridization. Starved animals were found to have four times fewer cells while maintaining cell abundance ratios, and differentially expressed genes were characterized across all cell types.

III. Conclusion

Whole-organism perturbation experiments can be read out using multiplexed scRNA-seq, even for relatively obscure non-model organisms. Analysis at the level of cell types gives a systems-level view of gene expression that will help elucidate coordinated gene regulation across entire animals.

References

[1] Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).

[2] Svensson, V., Vento-Tormo, R. & Teichmann, S. A. Exponential scaling of single-cell RNA-seq in the past decade. *Nat. Protoc.* **13**, 599–604 (2018).

[3] Han, X. *et al.* Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell* **172**, 1091-1107.e17 (2018).

[4] Stoeckius, M. *et al.* Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224 (2018).

[5] Kang, H. M. *et al.* Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.* **36**, 89–94 (2018).

[6] McGinnis, C. S. *et al.* MULTI-seq: Scalable sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *bioRxiv* (2018).

[7] Gehring, J., Park, J. H., Chen, S., Thomson, M. & Pachter, L. Highly Multiplexed Single-Cell RNA-seq for Defining Cell Population and Transcriptional Spaces. *bioRxiv* (2018).

[8] La Manno, G. *et al.* RNA velocity of single cells. *Nature* **560**, 494–498 (2018).

[9] Haghverdi, L., Büttner, M., Wolf, F. A., Buettner, F. & Theis, F. J. Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* **13**, 845–848 (2016)

[1]Department of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA. E-mail: jgehring@caltech.edu

[2]Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA.

A simple and generalizable metric for quantifying feedback control in biomolecular systems

Mariana Gómez-Schiavon[1], and Hana El-Samad[2]

Feedback control is ubiquitous in biomolecular networks. However, a metric to quantify biological feedback control and compare different control strategies do not exist. Here, we propose such a metric that derives the advantage of a feedback control system compared to a series of locally analogous systems without feedback. Each locally analogous system is constructed by substituting the feedback from the regulated process to the controller system by a constant input, such that its internal state is analogous to the original system. Thus, before a perturbation, each locally analogous system behaves identically to the feedback system. In this way, the impact of all the intrinsic biomolecular constraints of the system are being considered, while effectively isolating the effect of feedback control. We demonstrate that this metric can be applied to many feedback control systems, regardless of the underlying complexity of the biomolecular network, producing a readily interpretable value. Using this metric, we systematically analyze multiple feedback control biomolecular motifs proposed in the literature, and demonstrate that we can recapitulate established knowledge while providing novel insights. For example, high dilution rate is known to have a destructive effect on feedback control; the metric recognizes this behaviour, while also revealing that feedback control by active degradation motif can be optimized at intermediate dilution rates under biologically realistic parameters. In summary, we present a simple, generalizable and informative metric that allows to characterize feedback control in biomolecular contexts, and can guide our efforts on dissecting and designing biomolecular feedback control.

[1]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA. Email: Mariana.GomezSchiavon@ucsf.edu

[2]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA. Chan-Zuckerberg Biohub, USA. Email: Hana.El-Samad@ucsf.edu

Cooperative Control of Bacterial Consortia with Applications to a Wound-Healing Model System

Leopold N. Green[1], Chelsea Hu[1], Xinying Ren[2] and Richard M. Murray[1,2]

***Abstract*— Wound healing is a complicated biological network consisting of many types of cellular dynamics and functions regulated by chemical and molecular signals. Recent advancements in synthetic biology has made it possible to predictably design and build close-looped controllers that can function appropriately alongside of biological species. In this study, we developed a simple wound healing model mimicking the sequential relay-like dynamics of cellular populations involved in the healing process. We also developed a set of I1-FFL controllers that can sense the change in acute chronic wound dynamics, fixing the system in a timely manner.**

***Index Terms*— feed-foward loop controllers, population controllers, synthetic biology, wound healing**

I. Introduction

Wound healing is a dynamic, multi-cellular process regulated by a complicated network of propagating cell signals [1]. Healthy response to tissue injury relies on a systematic cascade of events known as acute wound healing. Acute healing is classically defined by four overlapping phases, distinct in function and histological characteristics: Hemostasis Phase - blood coagulation; Inflammatory Phase - wound debridement; Proliferation Phase -formation of granulation tissue and wound closure; and Remodelling Phase - improved wound tensile strength after closer [2]. Together, the phases demonstrate relay-like dynamics of cellular densities and functions, coordinated by the secretion of various signals [3].

Some of the prominent cells involved in wound healing are platelets, neutrophils, macrophages, and fibroblasts. During healing, cells enter the point of injury by either cell migration, infiltration, proliferation, and differential. These dynamics are controlled by an equally sophisticated signalling network of growth factors, cytokines and chemokines [4].

Platelets are the first cell type to enter the wound, releasing pro-inflammatory growth factors and cytokines promoting the migration of neutrophils to the site of injury. As the first circulating inflammatory cells to enter the site of injury, neutrophils secrete additional pro-inflammatory signals for macrophage migration into the wound.

The inflammation cycle transitions into the proliferation phage as macrophage-derived growth factors reaches optimal levels causing an influx of fibroblasts into the wound. Fibroblasts release extra-cellular matrix components promoting wound closure. Finally, during the remodelling phase, fibroblasts differentiate into contractile myofibroblasts resulting in minimal scar tissue and preserved tissue function.

The orchestrated interactions of various cell types, extracellular components, growth factors, and cytokines together play important roles in each of the different stages during healing. Therefore, an imbalance to any of these elements may lead to either prolonged, chronic healing.

II. Results

Our computational approach focuses on the elucidation of control systems needed to sense and regulate against impaired dynamics of complex signalling networks. We propose a multiple layer population controller consisting of a wound healing circuit that demonstrates cellular population dynamics of acute physiological wound healing (layer 1), and the coordination of feedback controllers that sense chronic dynamics, fixing the system back to acute healing (layer 2).

In this work, (1) we simulate a deterministic model of two coupled negative feedback modules producing sequential cellular population dynamics resembling the physiological acute wound healing process; (2) we implement predictive chronic wounds by altering parameters resembling physiological hyper-inflammation and impaired proliferation conditions; and (3) we regulate against various chronic conditions using two pulse generating type-1 incoherent feed-forward loops (I1-FFL) [5] coupled to the wound healing layer by signals produced by cellular species.

Our simulation results demonstrate that when chronic dynamics persists, coupling our pulse generating regulator controller circuits consisting of two I1-FFL modules is an effective design for regulating against incomplete healing. However, there are some conditions that our proposed I1-FFL controllers fail to regulate against. For more robust chronic regulation, future controller circuits may require a combination of multiple controller architectures and functions.

This work is supported by Defense Advanced Research Projects Agency (Agreement HR0011-17-2-0008).

[1]Department of Biology and Biological Engineering, California Institute of Technology. [2]Department of Control and Dynamical Systems, California Institute of Technology, Pasadena, CA.

References

[1] Schwartz, et al. Schwartz's Principles of Surgery. 9th ed., McGraw-Hill Education, 2014.

[2] Li, et al. Pathophysiology of acute wound healing. Clinics in dermatology. 2007.

[3] Sun, et al. Advances in skin grafting and treatment of cutaneous wounds. Science. 2014.

[4] Barrientos, et al. Growth factors and cytokines in wound healing. Wound repair and regeneration. 2008.

[5] Mangan S, Alon U. Structure and function of the feed-forward loop network motif. PNAS. 2003.

Random plasmid partitioning reduces the evolutionary stability of synthetic circuits

Andrew D. Halleran[1], Emanuel-Flores Bautista[2], Richard M. Murray[1,3]

***Short Abstract* — Plasmids are found across bacteria, archaea, and eukaryotes and play an important role in evolution. Plasmids exist at different copy numbers, ranging from a single plasmid per cell to hundreds. This feature of a copy number greater than one can lead to a population of plasmids within a single cell that are not identical clones of one another. During cell division, this population of plasmids is partitioned into two daughter cells, resulting in a random distribution of plasmid variants in each daughter. In this study, we use stochastic simulations to investigate how random plasmid partitioning impacts evolution of plasmid-encoded traits.**

***Keywords* — plasmid segregation, partitioning, evolution, synthetic biology**

I. Introduction

Plasmids occur naturally in bacteria, archaea, and eukaryotes [1-3]. Most commonly found in bacteria, plasmids are typically small, extra-chromosomal, stretches of DNA that replicate independently of the host genome. Plasmid copy number, the average number of plasmids per cell, can vary from ~1 copy per cell to hundreds depending on the mechanism regulating plasmid replication. Plasmids are a key source of genetic diversity and play an important role in evolution, and recent work has demonstrated the ability of high copy number plasmids to expedite acquisition of antibiotic resistance [4-5]. Plasmids are also widely used in synthetic biology and biotechnology, where the ability to rapidly engineer a plasmid and transform it into a host for maintenance and expression is of great value. Despite their widespread abundance in the natural world, and their frequent use in synthetic biology and biotechnology, the evolution of plasmid-encoded traits remains incompletely understood.

Plasmids have diverse inheritance modes [6-8]. Plasmids with high copy numbers typically rely on random binomial partitioning at cell division to ensure both daughter cells receive the plasmid [9]. This strategy is not viable for plasmids with low copy numbers as there is a high likelihood a daughter cell would not inherit the plasmid [10]. Low copy number plasmids instead use active partitioning systems, which distribute an equal number of plasmids to each daughter cell during division [8, 11-12].

Due to the independent nature of mutations, a given cell can contain a mix of plasmids where some contain the mutation and others do not. A similar phenomenon exists in mitochondria which can contain multiple distinct mitochondrial genomes per organelle [13]. Thus, for both low and high copy number plasmids a key feature of their inheritance is the random partitioning of a potentially mixed pool of alleles into daughter cells. Previous work has experimentally demonstrated random plasmid partitioning accelerates evolution, but little understanding exists about the parameter dependence (plasmid copy number, burden, mutation rate, and selection mode) of this effect [14-15]. Our work focuses on how random plasmid partitioning differs from perfect partitioning, and the impact of changing context-dependent rate parameters.

II. Results

Our simulation results demonstrate that random plasmid partitioning accelerates mutant allele fixation when the allele is beneficial and the selection is in an additive or recessive regime where increasing the copy number of the beneficial allele results in additional benefit for the host. This effect does not depend on the size of the benefit conferred, or the plasmid mutation rate, but is magnified by increasing plasmid copy number.

III. References

1. Couturier *et al.,* 1988 ; 2. Wang *et al.,* 2015 ; 3. Meinhardt *et al.,* 1990 ; 4. Harrison and Brockhurst, 2012 ; 5. San Millan *et al.*, 2017 ; 6. Novick, 1987 ;7. Summers, 1991 ; 8. Williams and Thomas, 1992 ; 9. Summers and Sherratt, 1984 ; 10. Peterson and Phillips, 2008 ; 11. Ogura and Hiraga, 1983 ; 12. Motallebi-Veshareh *et al.,* 1990 ; 13. Payne *et al.,* 2012 ; 14. Tyo *et al.,* 2009 ; 15. Ilhan *et al.,* 2018

1. Bioengineering, California Institute of Technology, Pasadena, CA, USA
2. Biology, Universidad Autonoma de Yucatan, Merida, Yucatan, Mexico
3. Control and Dynamical Systems, California Institute of Technology

Heritability of Fitness

Jungmin Han[1] and Vipul Periwal[2]

***Short Abstract*— By the definition of fitness, an organism is considered to be fit if its offspring are able to reproduce. We aim to investigate the heritability of this fitness through a series of iterative simulations that models the evolution of a single organism to a population, including mutation, damage repair, reproduction, and death. Our simulations showed that evolved networks possess heritable fitness that can be passed down from one generation to the next, suggesting the genetic robustness of these networks against mutation is an emergent trait.**

***Keywords*— gene network, evolution, fitness**

I. Purpose

Since the beginning of time, evolution has been acting on the traits of every biological organism. There have been many different approaches to explain the evolvability and genetic robustness of an organism [1, 2]. One of the key features of evolutionary theory is the definition of the "fitness" of an organism, or how well an organism can pass down its genetic traits to its following generations. For an organism to be fit, its genotype needs to be robust against mutations and environmental fluctuations. For the purpose of this study, we defined fitness to be the number of grandchildren.

We constructed a mathematical model that simulates the evolution of a genome to a large population over time. We evaluated how fitness evolves within the population. Furthermore, the heritability of fitness from one generation to the next was examined. The model simulations showed that there is a positive correlation between the fitness of a parent and that of its offspring.

II. Methods

A $n \times n$ matrix W, representing a gene network, and a $1 \times n$ vector x, representing an initial phenotypic state, were assigned to the founder organism. The dynamics of the phenotypic state was determined from a Poisson distribution:

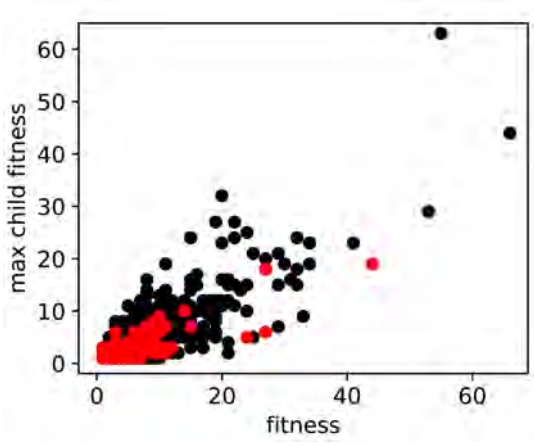
$$x(t+1) \sim P\left(x(t)(1-\rho) + \rho\sqrt{max(0, x \cdot W + \eta)}\right), \quad (1)$$

where ρ is a decay factor and η is noise. The genome of an organism at time t, denoted as y, was a specific linear combination of its phenotypic state $x(t)$. At any time point, if an organism satisfies $y(t) \geq 2y(0)$, indicating genome replication, it generates an offspring. The offspring inherited a mutated gene network from its parent, and its phenotypic state, x', was approximately half of the parent's, while the parent was left with the other half of the state. This mimics the partitioning of the contents of a cell during cell division.

The gene network of every organism undergoes somatic mutations during its life. Organisms are able to expend simulation steps to repair their gene network, if it was too far off from the initial W. Finally, we randomly pruned the population every now and then to simulate a finite carrying capacity of the environment.

III. Conclusion

We simulated the model over some iterations to generate a large population and studied the relationships between parents' fitness and key characteristics of their offsprings' fitness. At the end of the simulation, we analyzed the structure of the tree describing the population's evolution. The scatter plot of the parents' fitness vs. the maximum of their offsprings' fitness is shown below. Red dots represent active organisms that are still able to reproduce, while the black dots are the organisms that were randomly removed from the population during the pruning process. The figure shows a positive correlation between the two variables, suggesting that an organism's ability to reproduce was preserved from generation to generation, indicating genetic robustness is an emergent trait.



Acknowledgments: This work was funded by the Intramural Research Program of the NIH, NIDDK.

[1]Laboratory of Biological Modeling, National Institute of Diabetes and Digestive and Kidney Diseases, NIH. Email: jungmin.han@nih.gov

[2]Laboratory of Biological Modeling, National Institute of Diabetes and Digestive and Kidney Diseases, NIH. Email: vipulp@niddk.nih.gov

References

[1] K. Kaneko. Evolution of robustness to noise and mutation in gene expression dynamics. *PLoS One*, 5:e434, 2007.

[2] M. L. Siegal and A. Bergman. Waddington's canalization revisited: Developmental stability and evolution. *PNAS*, 99(16):10528–10532, 2002.

Interaction Classifier Improves Accuracy of Fully Automated Model Expansion

Casey Hansen[1], Peter Spirtes[2], Cheryl Telmer[3], Kara Bocan[4], Emilee Holtzapple[5], Yasmine Ahmed[4], Adam Butchy[1], Khaled Sayed[4], Natasa Miskov-Zivanov[1,4,5]

***Short Abstract* — We present an Interaction Classifier (IC), a tool for comparing, sorting, and ranked scoring of biological interactions automatically extracted from published literature. The IC determines a score for each interaction to quantify their relevance and usefulness. The sorting and scoring significantly reduce the amount of expert intervention required to evaluate the importance and relevance of a large number of interactions from diverse literature sources. The IC also performs a classification of interactions into corroborations, contradictions, and extensions, crucial for validating and expanding models. The utility of the IC is demonstrated with data sets containing differing types of knowledge.**

***Keywords* — Natural Language Processing, Machine reading, Text mining, Automated model generation**

I. Introduction

Computational models are an excellent means for running *in silico* experiments to understand biological phenomena and test or guide new hypotheses. These models can be created with expert knowledge or using data and interactions found in the literature. Using information from the literature allows the inclusion of results from many experts and experiments, from diverse disciplines, however, it can be time-consuming to manually investigate every relevant article for applicable knowledge. Natural Language Processing (NLP) tools such as REACH [1] and TRIPS [2] can perform this task automatically, extracting, and standardizing information from a large number of articles.

NLP tools can collect large amounts of processed text with relative ease, but expert opinion is still required to determine what extracted knowledge is most relevant to the model of interest. Some information extraction approaches, like those used by RUBICON [3,4] and INDRA [5], include a believability score, as a way to measure the quality of the extraction. However, due to the variability of language in research literature, a meaningful belief score can be difficult to define and compute.

II. Interaction Classifier

Here, we present an Interaction Classifier (IC), a tool that compares an existing model, and the interactions it is composed of, to those extracted automatically from literature. Interactions are first classified as corroborations, contradictions, or extensions with respect to the model, and then within these broader classifications they are given a numerical score based on a more detailed comparison of the attributes of the interaction. From this, we create a quantitative measure of interaction *relevance*, how well the interaction matches the model, and interaction *usefulness*, how much information it can add to the model. These measures significantly reduce the burden of manual inspection of thousands of interactions in thousands of papers.

We demonstrate an application of the IC with a previously-created model of a melanoma cell line, and we evaluate the IC scoring by comparing it to expert judgment. Our analysis included three sets of literature extractions: a large set compiled using relevant search terms from REACH, a smaller set created using a more directed search, and a final set using expert judgement to determine relevance and usefulness of the compiled interactions from the directed search. We used the IC to compare the model to each set of interactions, and then manually inspected the scoring outcomes of those interactions that were classified as extensions by the IC. We compared the frequency of extraction and scores of the interactions from the judged reading to those interactions from the larger reader sets.

III. Results and Conclusion

The results of this study provide scoring metrics and guidelines for which scores and attributes can be used to automatically identify relevant and useful interactions for evaluating and improving a computational model. We found two important characteristics in identifying the highest quality interactions: the number of times an interaction is extracted from the relevant literature (frequency), and the details of the extracted interaction that are not already in the corresponding model interaction (usefulness).

Acknowledgements: This work was partially supported by DARPA award W911NF-17-1-0135.

[1]Department of Bioengineering, University of Pittsburgh, USA. E-mail: {ceh92,aab133}@pitt.edu

[2]Department of Philosophy, Carnegie Mellon University, USA. E-mail: ps7z@andrew.cmu.edu

[3]Department of Biological Sciences, Carnegie Mellon University, USA. E-mail: ctelmer@cmu.com

[4]Department of Computational Biology, University of Pittsburgh, USA. E-mail: erh87@pitt.edu

[5]Department of Electrical and Computer Engineering, University of Pittsburgh, USA. E-mail: {knb12,yaa38,k.sayed,nmzivanov}@pitt.edu

References

[1] REACH website, http://agathon.sista.arizona.edu:8080/odinweb/

[2] Ferguson G, Allen JF. "TRIPS: An Integrated Intelligent Problem-Solving Assistant," Proceeding of the Fifteenth National Conference on AI, Madison, WI 1998

[3] Burns GA, et al., "Automated detection of discourse segment and experimental types from the text of cancer pathway results sections." In: Database 2016, p. baw122 (2016)

[4] Sayed K, Telmer CA, Butchy AA, Miskov-Zivanov N, "Recipes for Translating Big Data Machine Reading to Executable Cellular Signaling Models." In: Machine Learning, Optimization, and Big Data (MOD), Volterra, Italy, September 2017.

[5] INDRA website, https://indra.readthedocs.io/en/latest/index.html

A computational model of cell phenotype interactions in small cell lung cancer

Leonard Harris[1], Samantha Beik[2], Sarah Groves[3], Aliss

***Abstract*— Small cell lung cancer (SCLC) is an aggressive neuroendocrine carcinoma known for rapid recurrence and metastasis following treatment. While histologically homogeneous, recent work suggests that SCLC tumors are comprised of numerous subtypes that support growth and facilitate treatment avoidance and survival. Here, we use computational modeling to investigate phenotype interactions in SCLC tumors. Specifically, we generate multiple population dynamics models of phenotypic interactions in SCLC and fit to human and mouse tumor data to identify driving factors differentiating tumor types. Our results indicate that cell-cell interactions are crucial for maintaining stable phenotypic content and suggest experimentally-testable interventions for modulating tumor composition.**

I. Background

SMALL cell lung cancer (SCLC), the most lethal form of lung cancer (5-10% 1-yr survival), is characterized by rapid growth, early metastasis, and treatment-resistant relapse after initial positive response. The standard of care (etoposide + platinum-based chemotherapeutics) has not changed in ∼30 years, emphasizing the need for new and improved therapeutic approaches. Though traditionally thought of as a homogeneous disease comprised of "small round blue cells," recent research has identified a robust tumor microenvironment comprised of numerous phenotypic variants with synergistic functions and interactions [1–4].

Recently, we identified four distinct SCLC subtypes using consensus clustering and weighted gene co-expression network analysis (WGCNA) on transcriptomics data from over 50 patient-derived cell lines [4]. Three of these subtypes had been described previously [1–3], while the fourth shows broad insensitivity to several classes of therapeutic agents, suggesting a possible role in treatment resistance and tumor recurrence. Subsequent analysis using CIBERSORT [5] on human and mouse tumor samples indicated that SCLC tumors are comprised of varying proportions of all four subtypes [4]. Here, we use population dynamics modeling and Monte Carlo-based parameter fitting to identify factors, such as differentiation rates and interaction modes/strengths, that distinguish these tumor types. Our insights provide potential avenues for modulating the phenotypic compositions of SCLC tumors, which can be tested experimentally.

[1]Department of Biochemistry, [2]Cancer Biology Graduate Program, [3]Chemical and Physical Biology Graduate Program, [4]Department of Cell and Developmental Biology, [5]Department of Pathology, Microbiology, and Immunology, [6]Department of Pharmacology, and [7]Department of Biomedical Informatics, Vanderbilt University, Nashville, TN, USA.

[8]Email: vito.quaranta@vanderbilt.edu

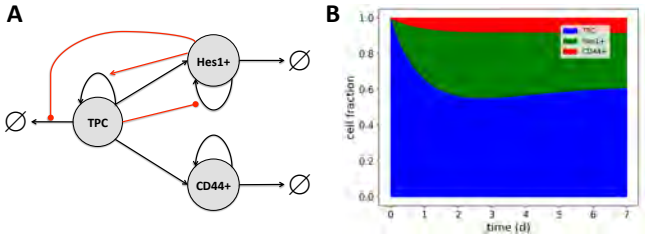


Fig. 1. **Preliminary model of phenotype interactions in SCLC.** (A) Tumor propagating cells (TPCs) can differentiate into two phenotypes; all phenotypes can divide and die; and some can enhance or inhibit division and/or death of others (shown in red). (B) Interactions between phenotypes leads to stabilization of the phenotypic composition over time.

II. Results

We began by considering three SCLC phenotypes [2]: tumor-propagating cells (TPCs), Hes1+ cells, and CD44+ cells (Fig. 1A). The model includes differentiation of TPCs into both Hes1+ and CD44+ phenotypes and interactions between TPCs and Hes1+ cells through secreted factors that affect division and/or death rates (Fig. 1A). Experimental data indicate that *in vivo* SCLC tumors are composed of ∼50% TPCs, 25% Hes1+, 5% CD44+, and 20% unknown cell types. By adjusting the rate constants of the model we were able to reproduce this observation *in silico* (Fig. 1B). Specifically, the rate of differentiation from TPCs to Hes1+ cells had to be approximately 5× larger than for TPCs to CD44+, providing a testable prediction of the model.

We are in the process of reformulating and extending this model within the current view of SCLC tumors as comprised of four phenotypes [4]. This includes accounting for trophic factors such as midkine and fibroblast growth factor (FGF) and fitting across human and mouse tumors with varying phenotypic composition [4]. Preliminary results confirm that differentiation rates are a major determinant of tumor composition as are feedbacks among phenotypes.

References

[1] Borromeo, M. D. *et al.* ASCL1 and NEUROD1 reveal heterogeneity in pulmonary neuroendocrine tumors and regulate distinct genetic programs. *Cell Rep.* **16**, 1259–1272 (2016).

[2] Lim, J. S. *et al.* Intratumoural heterogeneity generated by Notch signalling promotes small-cell lung cancer. *Nature* **545**, 360–364 (2017).

[3] Mollaoglu, G. *et al.* MYC drives progression of small cell lung cancer to a variant neuroendocrine subtype with vulnerability to aurora kinase inhibition. *Cancer Cell* **31**, 270–285 (2017).

[4] Wooten, D. J. *et al.* Systems-level network modeling of small cell lung cancer subtypes identifies master regulators and destabilizers. *bioRxiv* (2019). https://doi.org/10.1101/506402.

[5] Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453–457 (2015).

Protein contacts using Expectation Reflection

Danh-Tai Hoang, Joseph McKenna, Chris Yang, and Vipul Periwal

***Short Abstract* — We recently developed a data-driven method, Expectation Reflection, for network inference in stochastic systems. This method outperforms present state-of-the-art methods in predicting the interaction weights between binary variables. Here, we extend the method to recover interaction networks of categorical variables. Applying this to protein multiple sequence alignments, we infer the residue-residue interactions in protein domains. We then calculate the direct information and predict the correlations between positions in the protein sequences. We find good prediction results for many protein domains.**

***Keywords* — Network reconstruction, protein structure prediction, protein multiple sequence alignments.**

I. Purpose

PROTEIN function is greatly dependent on the arrangement of amino acids in a three-dimensional structure. Therefore, the determination of tertiary structures of proteins from their primary sequences is a fundamental problem in biology. Direct Coupling Analysis (DCA) is the present state-of-the-art method for protein structure inference. This method identifies interactions between residues at different positions based on phylogenetic correlations between the residues [1].

We recently developed an approach, Expectation Reflection (ER), and demonstrated that ER works better than existing methods in predicting coupling weights between binary variables, especially in the limit of small samples sizes and partially observed systems [2, 3].

Here, we extend ER to infer networks of interactions from observed configurations of (non-binary) categorical variables such as protein sequences.

II. METHOD

We applied a one-hot encoding to represent categorical variables. We used an iterative algorithm to update the model of the influence on each position from other positions according to the ratio of the observation to the corresponding model expectation. This approach completely separates model updates from minimization of a cost function measuring goodness of fit, so that this cost function can be used as the stopping criterion of the iteration. This is crucial for biological applications where available data usually under-determines complex systems.

Acknowledgements: This work was supported by the Intramural Research Program of the National Institutes of Health, NIDDK (D.-T.H., V.P.).

Laboratory of Biological Modeling, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.
Email: danh-tai.hoang@nih.gov, vipulp@mail.nih.gov

III. RESULTS

We first tested our method to infer the connection strengths from synthesized sequences from simulated stochastic models of coupled categorical variables. We then applied our method to multiple sequence alignments of protein families from the Pfam database [4], inferring interaction strengths between residues in protein domains. From these interactions, we calculated the direct information [1] determining correlations between different positions in a protein chain. The predictive performance was validated by comparing to structures obtained from PDB. We computed the receiver operating characteristic from our method and compared with DCA (Figure 1).

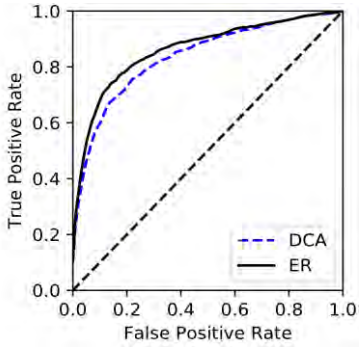


Figure 1. Receiver operating characteristic (ROC) curve from a set of 15 protein domains, shown for ER (black) and DCA (blue).

IV. Conclusion

Extending our ER method, we propose an effective algorithm of network inference for categorical variables. Applying this method to multiple sequence alignments of protein families, we can improve the inference of contacts between residues in 3D structures of proteins.

References

[1] F. Morcos, et al. (2011) Direct-coupling analysis of residue coevolution captures contacts across many protein families, PNAS, **108** (49) E1293-E1301.

[2] D. T. Hoang, J. Song, V. Periwal, and J. Jo (2019), Network inference in stochastic systems from neurons to currencies: Improved performance at small sample size, Physical Review E, **99**, 023311.

[3] D. T. Hoang, J. Jo, and V. Periwal (2019), Data-driven inference of hidden nodes in networks, arXiv:1901.04122.

[4] R. D. Finn, et al. (2015), The Pfam protein families database: towards a more sustainable future, Nucleic Acids Res, **44**, D279-D285.

Repetitive sequences tune transcription factor-DNA binding by modulating the kinetics of the transcription factor search process

Connor A. Horton[1] and Polly M. Fordyce[1,2,3,4,*]

***Short Abstract* — Transcription factors (TFs) regulate gene expression via sequence-specific binding to regulatory elements within genomic DNA. Despite the enormity of the potential genomic search space, TFs rapidly locate and bind their cognate sites by combining three-dimensional diffusion with one-dimensional sliding, hopping, and intersegmental transfer. We apply a microfluidic platform to measure the equilibrium binding affinities and kinetic parameters for two transcription factors interacting with a library of >30 DNA sequences. Our results suggest that regulatory sequences flanking known TF consensus sites might provide a cellular mechanism to 'tune' TF binding to continuously modulate levels and kinetics of transcription *in vivo*.**

***Keywords* — protein-DNA binding, transcription factor binding, transcriptional regulation, microfluidics**

Transcription factors (TFs) regulate gene expression via sequence-specific binding to regulatory elements within genomic DNA. Despite the enormity of the potential genomic search space, TFs rapidly locate and bind their cognate sites by combining three-dimensional diffusion with one-dimensional sliding, hopping, and intersegmental transfer [1-3]. Previous experimental measurements have demonstrated that repetitive sequences surrounding TF cognate sites significantly increase relative levels of bound TFs [4]; however, the effects of repetitive sequences on the equilibrium levels and kinetics of TF binding have not been directly measured in high-throughput.

To address this, we have applied a microfluidic platform (MITOMI, or Mechanically-Induced Trapping of Molecular Interactions), to measure the equilibrium binding affinities (K_d) and association and dissociation rates (k_{on}, k_{off}) for different transcription factors interacting with a library of >30 DNA sequences [5]. This library includes sequences containing the known Pho4 and MAX consensus site (5'-GTCACGTGAC-3'), a mutated consensus site (5'-GTCACGCGAC-3'), and a shuffled consensus site, each flanked by various repetitive or random flanking sequences. We have assayed two homodimeric basic-helix-loop-helix transcription factors—Pho4, a TF from *S. cerevisiae* that mediates the transcriptional response to phosphate starvation, and Max, a human TF implicated in several oncogenic pathways. We provide additional evidence that TFs consistently bind to repetitive DNA sequences with higher affinity than to random sequences. We also establish that the difference in measured K_d values are primarily driven by changes in k_{on}. This effect persists even in the absence of Max and Pho4's cognate binding site (the E-box, or CACGTG), supporting the model that nonspecific binding to repetitive sequences occurs as a distinct mechanism from binding to the cognate site.

Together, these results suggest that regulatory sequences flanking known TF consensus sites might provide a cellular mechanism to 'tune' TF binding to continuously modulate levels and kinetics of transcription *in vivo*. In current work, we are constructing synthetic Pho4-driven promoters to directly measure *in vivo* effects of repetitive sequences on transcriptional dynamics. We are also probing the mechanism by which these nonspecific sequences tune binding kinetics by systematically measuring the binding affinities and kinetics of hundreds of Pho4 mutants to DNA sequences with varying degrees of repeats. These high-throughput and quantitative measurements of biophysical parameters will further elucidate the TF search process and contribute to a greater theoretical understanding of gene regulation.

References

[1] Berg OG, Winter RB, von Hippel PH (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* **20**, 6929–48.

[2] Halford SE, Marko JF (2004). How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res.* **32**, 3040–52.

[3] Gorman J, Greene EC (2008). Visualizing one-dimensional diffusion of proteins along DNA. *Nat. Struct. Mol. Biol.* **15**, 768–74.

[4] Afek A, et al. (2014). Protein-DNA binding in the absence of specific base-pair recognition. *Proc. Natl. Acad. Sci.* **111**, 17140–45.

[5] Fordyce PM, et al. (2010). De novo identification and biophysical characterization of transcription-factor binding sites with microfluidic affinity analysis. *Nat. Biotechnol.* **28**, 970–75.

[1]Department of Genetics, Stanford University, Stanford, CA 94305.
[2]Department of Bioengineering, Stanford University, Stanford, CA 94305
[3]ChEM-H Fellow, Stanford University, Stanford, CA 94305.
[4]Chan Zuckerberg Biohub, San Francisco, CA 94158.
*Correspondence to pfordyce@stanford.edu

Acetaminophen-induced liver injury: interhepatocyte communication

Ryan C. Kennedy[1], Andrew K. Smith[1], Glen E.P. Ropella[2], Mitchell R. McGill[3], Hartmut Jaeschke[4], and C. Anthony Hunt[1]

***Short Abstract* — Previous virtual model mechanisms provided the first concrete multi-attribute explanation for how and why liver necrosis begins near the central vein (CV). Such mechanisms lacked the characteristic feature that subsequent necrosis occurs primarily adjacent to hepatocytes that have already initiated necrosis. To facilitate, we added logic enabling hepatocytes to obtain information about neighboring hepatocytes. So doing allowed virtual hepatocytes to achieve new target attributes. Additional logic for extracellular vesicle (EV) objects was added to provide an alternative explanation. We hypothesize that the resulting intercellular-communication-enabled model mechanism is analogous to the actual explanation for APAP-induced hepatotoxicity at comparable levels of granularity.**

I. Introduction

ACETAMINOPHEN (APAP) induced liver injury is clinically significant, and APAP overdose in mice often serves as a model for drug-induced liver injury in humans. Previous studies provided an explanation for how and why necrosis first begins near the CV before propagating [1]. We add new target attributes and virtual experiments in which we 1) require necrosis events to first occur adjacent to the CV and 2) specify that subsequent necrosis events must occur primarily adjacent to virtual hepatocytes (vHPCs) that have already triggered necrosis. Evidence supports that intracellular communication via exosomes [2], gap junctions [3,4], and connexin hemichannels [5] plays a vital role in determining the toxic effect of chemicals. Here, we add communication-enabled rule-based capabilities to facilitate meeting the target attributes.

II. Methods

We experiment on Control and Test Mice. Control Mice utilize validated model mechanisms and have been previously described [1]. We sought to achieve the new target attributes in Test Mice. Through virtual experiments, incremental changes to model mechanism features were made with the goal of improving the mechanism-based explanation.

A pair of rules were added to determine whether vHPCs in Test Mice initiate necrosis. For stressed vHPCs, necrosis is triggered if the vHPC is adjacent to the CV or if *n* or more downstream neighbors have initiated necrosis. For non-stressed vHPCs, *m* or more lateral neighbors must have initiated necrosis for the vHPC to initiate necrosis. Stressed vHPCs are defined as those that have accumulated a specified amount of mitochondrial damage. An additional mechanism involving the passing of EV objects was also explored.

Credibility of Test Mice is increased through validation and falsification. Virtual Mice are hypotheses for corresponding mouse features, and Test Mice are instantiated hypotheses. Hypothesis testing requires consideration of similarity of a Test Mouse to wet-lab experiments and similarity of variability within and between virtual and wet-lab experiments. Further details on the Virtual Mice and related information have been published [6].

III. Results and Discussion

Control Mice provide a plausible and coarse-grain explanation of early features of APAP-induced injury. vAPAP measurements from Test Mice and from Control Mice were indistinguishable and location-dependent per-vHPC amounts of various products were comparable. There are differences in the moving averages for the mean distance from the CV for necrosis events in Control and Test Mice. In Control Mice, necrosis events are triggered upstream of the CV, while in Test Mice, the averages start at the CV and creep upstream, while in Control Mice, necrosis events are triggered upstream of the CV. In achieving our target attributes, we can argue that analogous processes may occur in vivo. Literature supporting evidence that exosome-based communication is strong, and future experiments will focus on this aspect.

References

[1] Smith AK, et al. (2016). Competing mechanistic hypotheses of acetaminophen-induced hepatotoxicity challenged by virtual experiments, *PLoS Comput. Biol.* **12**(12), e1005253.

[2] Hirsova P, et al. (2016). Extracellular vesicles in liver pathobiology: Small particles with big impact. *Hepatology*. **64**(6), 2219-2233.Lastone AB, Lasttwo AB (1993) *Book style.* Wadsworth, Belmont, CA.

[3] Asamoto M, et al. (2004). Connexin 32 dominant-negative mutant transgenic rats are resistant to hepatic damage by chemicals. *Hepatology*. **40**(1), 205-210.

[4] Decrock E, et al. (2009). Connexin-related signaling in cell death: to live or let die? *Cell Death Differ*. **16**(4), 524.

[5] Maes M, et al. (2017). Connexin hemichannel inhibition reduces acetaminophen-induced liver injury in mice. *Toxicol. Lett*. **278**, 30-37.

[6] Kennedy RC, et al. (2019). Propagation of pericentral necrosis during acetaminophen-induced liver injury: evidence for early interhepatocyte communication and information-exchange, *Toxicol Sci.* kfz029, https://doi.org/10.1093/toxsci/kfz029

[1]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA. E-mail: ryan.c.kennedy@alumni.nd.edu; dr.andrew.k.smith@gmail.com; a.hunt@ucsf.edu

[2]Tempus Dictum, Inc., Milwaukie, OR. E-mail: gepr@tempusdictum.com

[3]Department of Environmental and Occupational Health, Fay W. Boozman College of Public Health, University of Arkansas for Medical Sciences, Little Rock, AR. E-mail: MMcgill@uams.edu

[4]Department of Pharmacology, Toxicology and Therapeutics, University of Kansas Medical Center, Kansas City, KS. Email : hjaeschke@kumc.edu

Nothing should be here on page 2! Please limit your abstract to a single page, and create a one-page .pdf file for submission.

A causal model mechanism designed for clinical translation explains elevated serum ALT in mice following toxic doses of Acetaminophen

Andrew K. Smith[1], Glen E.P. Ropella[2], Mitchell R. McGill[3], Hartmut Jaeschke[4], and C. Anthony Hunt[1]

***Short Abstract* — Previous virtual model mechanisms provided the first concrete multi-attribute explanation for how and why acetaminophen (APAP) induced liver necrosis in mice begins near the central vein. Elevated serum alanine aminotransferase (ALT) is a biomarker for such liver damage. The actual causal events linking APAP's reactive metabolite to elevated serum ALT are, however, still poorly understood. We used virtual experiments to concretize four virtual mechanism variants involving mitochondrial and non-mitochondrial damage products, and necrosis. They were able to map quantitatively to multiple wet-lab measurements in mice. Evidence supports one model mechanism as being most analogous to the actual explanation for elevated serum ALT at comparable levels of granularity.**

I. Objective

ACETAMINOPHEN (APAP) hepatotoxicity in mice is an established model for drug-induced liver injury in humans. Cell death and membrane damage are thought to be the most common causes of elevated serum alanine aminotransferase (ALT), which is a commonly used biomarker in clinical diagnosis and research involving liver injury induced by drugs and drug combinations [1]. However, the actual causal mechanisms linking APAP hepatic disposition and metabolism (including formation of the reactive metabolite NAPQI) to elevated serum ALT are still poorly understood. We use virtual experiments to investigate plausible explanatory multiscale model mechanisms [2] that may explain elevated serum ALT and select those that are both parsimonious and capable of achieving multiple quantitative validation targets.

II. Methods

We reuse the model mechanism developed by Smith et al. [3]. It is a concretized hypotheses for how features of APAP hepatotoxicity in mice are generated. It provides plausible quantitative causal explanations for major features of APAP hepatotoxicity in mice. We utilize the virtual experiment protocol outlined by Kirschner et al. [4], as updated [3]. Four requirements guide software engineering, mechanism instantiation, and simulation refinements. 1. Virtual entities are concrete and biomimetic in prespecified ways. 2. Entities and activities organized and orchestrated so that they are responsible for targeted features of hepatotoxicity. 3. Feature measurements match or are strongly analogous to Targeted Attributes [5] and meet prespecified quantitative Similarity Criteria. 4. Higher-level phenomena arise from component interactions at a lower-level.

Key features: NAPQI formation increases periportal (PP)-to-pericentral (PC) within hepatic lobules; probabilities of GSH depletion and damage mitigation events decrease PP-to-PC; in vivo damage products are represented by virtual mitochondrial (mitoD) and non-mitochondrial (nonMD) damage products. Results from studies in mice [5] provided quantitative validation targets.

III. Results and Discussion

We extend the above features by implementing model ALT release mechanism variants in which all virtual hepatocytes (vHPCs) initially contain the same amount of ALT. Stress resulting from accumulation of membrane damage above a threshold value triggers ALT release. When a vHPC becomes necrotic, all remaining ALT is released. We experimented on four mechanism variants: membrane damage is triggered by either accumulation of 1) mitoD, 2) nonMD, 3) the combination, or 4) only necrotic cells. Released ALT, which accumulates in virtual Mouse Body, maps to serum ALT. A parsimonious ALT release mechanism based on damage triggered by mitoD plus necrosis provide acceptable matches to validation targets and may be most analogous to the actual explanation for elevated serum ALT.

References

[1] McGill MR (2016) The past and present of serum aminotransferases and the future of liver injury biomarkers. *EXCLI journal* **15**, 817-828.

[2] Hunt CA, Erdemir A, Lytton W, Mac Gabhann F, et al. (2018) The spectrum of mechanism-oriented models and methods for explanations of biological phenomena. *Processes* **6**:56.

[3] Smith AK, Petersen BK, Ropella GE, Kennedy RC, Kaplowitz N, Ookhtens M, Hunt CA (2016) Competing mechanistic hypotheses of acetaminophen-induced hepatotoxicity challenged by virtual experiments. *PLoS Comput Biol.* **12**, e1005253.

[4] Kirschner DE, Hunt CA, Marino S, Fallahi-Sichani M, Linderman JJ (2014) Tuneable resolution as a systems biology approach for multi-scale, multicompartment computational models. Wiley Interdisci. Rev Syst Biol Med. 6, 289–309.

[5] McGill MR, Lebofsky M, Norris HR, Slawson MH, Bajt ML, Xie Y, Williams CD, Wilkins DG, Rollins DE, Jaeschke H (2013) Plasma and liver acetaminophen-protein adduct levels in mice after acetaminophen treatment: dose–response, mechanisms, and clinical implications. *Tox Appl Pharmacol* **269**, 240-249.

[1] Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA. E-mail: dr.andrew.k.smith@gmail.com; a.hunt@ucsf.edu

[2] Tempus Dictum, Inc., Milwaukie, OR. E-mail: gepr@tempusdictum.com

[3] Department of Environmental and Occupational Health, Fay W. Boozman College of Public Health, University of Arkansas for Medical Sciences, Little Rock, AR. E-mail: MMcgill@uams.edu

[4] Department of Pharmacology, Toxicology and Therapeutics, University of Kansas Medical Center, Kansas City, KS. E-mail: hjaeschke@kumc.edu

Modeling Emergent Biological Phenomena: A Case Study in the Social Amoeba

Chuqiao Huyan[1,2], Alexander Golden[2,3], Pankaj Mehta[2,3], and Allyson E. Sgro[1,2,3]

***Short Abstract* — Cells in groups display emergent biological phenomena. A widely studied example, *Dictyostelium discoideum*, the social amoeba, displays collective oscillations and movement coordinated by cAMP signal relay upon starvation. Past theoretical work has explored how these population-wide oscillations emerge through different models. Recent experimental advances have created a wealth of data about the key phenomena cells exhibit that drive these behaviors. This study validates the key experimental features recapitulated by different theoretical models. We identify each models' advantages and limitations, and describe behaviors in mutants that will allow us to identify how multicellular systems actively modulate these population-level emergent phenomena.**

EXTENDED ABSTRACT

GROUP oscillations spontaneously emerge in coupled cells when the environmental parameters like nutrient and cell density reach a critical value. Such examples include glycolytic oscillations in yeast [1], metabolite-dependent oscillations within biofilms [2], and oscillations in bacteria population carrying a synthetic gene circuit [3].

An extensively studied example is *Dictyostelium discoideum*, the social amoeba. Upon starvation, *Dictyosteium* cells initiate a developmental program, releasing cyclic-AMP (cAMP) in pulses. These pulses cause cells to produce their own cAMP which propagates the signal; at the same time, cAMP act as a chemoattractant that directs cell migration during aggregation. Both the cAMP signal relay and chemotaxis are critical for *Dictyostelium* cells to transition from unicellular to multicellular mound state [4]. Over the past few decades, multiple models have been developed aiming to describe this collective behavior based on cAMP signal relay. However, a comprehensive analysis of the existing models and how well they reproduce experimental observations is lacking.

This case study aims to compare and evaluate the existing models. Key phenomena that have been experimentally observed are summarized and models are evaluated by how well they recapitulate these qualitative and quantitative features validated by experiments. The models we evaluate include 1) models based on kinetic equations derived from molecular networks [5,6,7], 2) models that treat cells as simple automata [8,9], 3) a phase equations-based model that describes the onset of collective oscillations [10], and 4) models that capture key input-output responses ignoring detailed molecular networks [11,12]. The key experimental features evaluated in this work are 1) single cell responses including signal relay, adaptation, and super-threshold oscillations, and 2) population behaviors including synchronized / unsynchronized oscillations and spatial pattern formation. Taking advantage of this case study, we seek a theoretical interpretation of aberrant behaviors in mutant strains. A PKA regulatory element mutant develops rapidly after starvation and displays aberrant morphogenesis at the later developmental stage [13]. To elucidate the source of this aberrant behavior in this PKA mutant, we harness recently acquired quantitative measurements of single-cell level signaling dynamics with a genetically-encoded sensor. Considering the key features we find in different models, we propose that the mutant behavior can be effectively described by a top-down model [11], but in a bi-stable regime distinct from this published model.

REFERENCES

[1] Chandra, F. A., Buzi, G., & Doyle, J. C. (2011). Science, 333(6039), 187-192.
[2] Liu, Jintao, et al. (2015) "Metabolic co-dependence gives rise to collective oscillations within biofilms." *Nature* 523.7562: 550.
[3] Danino, T., Mondragón-Palomino, O., Tsimring, L., & Hasty, J. (2010). *Nature*, 463(7279), 326.
[4] Durston, A. J. (2013). *Current genomics* 14.6: 355-360.
[5] Martiel, Jean-Louis, and Albert Goldbeter. (1987). *Biophysical journal* 52.5: 807-828.
[6] Tyson, J. J., Alexander, K. A., Manoranjan, V. S., & Murray, J. D. (1989). *Physica D: Nonlinear Phenomena*, 34(1-2), 193-207.
[7] Maeda, Mineko, et al. (2004) "Periodic signaling controlled by an oscillatory circuit that includes protein kinases ERK2 and PKA." *Science* 304.5672: 875-878.
[8] Levine, H., Aranson, I., Tsimring, L., & Truong, T. V. (1996). *PNAS*, 93(13), 6382-6386.
[9] Sawai, S., Thomason, P. A., & Cox, E. C. (2005). *Nature*, 433(7023), 323.
[10] Gregor, T., Fujimoto, K., Masaki, N., & Sawai, S. (2010). The onset of collective behavior in social amoebae. *Science*, 328(5981), 1021-1025.
[11] Sgro, Allyson E., et al. (2015) "From intracellular signaling to population oscillations: bridging size‐ and time‐scales in collective behavior." *Molecular systems biology* 11.1: 779.
[12] Kamino, Keita, et al (2017) "Fold-change detection and scale invariance of cell–cell signaling in social amoeba." *PNAS* 114.21: E4149-E4157.
[13] Zhang, Hui, et al. (2003) "Constitutively active protein kinase A disrupts motility and chemotaxis in Dictyostelium discoideum." Eukaryotic Cell 2.1: 62-75.

Acknowledgements: This work was supported by a Burroughs Wellcome Fund Career Award at the Scientific Interface to A.E.S. and NIH grant R35GM119461 to P.M.

[1] Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA
[2] Biological Design Center, Boston University, Boston, MA 02215, USA
[3] Department of Physics, Boston University, Boston, MA 02215, USA

Construction of AND logic gate synthetic circuit using ribosome recovery system in *E. coli*

Prajakta K. Jadhav[1] and Nicholas C. Butzin[1]

***Short Abstract* — Orthogonal logic gates are required for construction of more complex systems in synthetic biology. Here we have developed an orthogonal AND logic gate in *E. coli* that has both transcriptional and translational control and exploits the natural ribosome recovery system. Ribosome rescue occurs when ribosomes stall, which can occur in the presence of rare codons (limited available cognate tRNA). Our system consists of consecutive rare codons, which allows us to regulate output using inducible promoters and by altering cognate rare tRNA expression. Our goal is to develop robust orthogonal logic gates as building blocks for complex circuits.**

***Keywords* — Ribosome stalling, ribosome rescue rare codon, AND logic gate**

I. Full Abstract

SYNTHETIC biology is an emerging field that combines biology and engineering, which allows scientists to introduce controllable, synthetically designed biological systems. These systems include oscillators, logic gates, toggle switches, pulse generators, biosensors, memory elements, and many more[1-4]. The design and implementation of these systems depend on developing biological building blocks for increased complexity. Currently, simple logic gates (AND, OR, NOR, etc.) in biological systems are gaining attention as they can act as interchangeable parts[5]. One of the biggest challenges for synthetic biologist is to build well-regulated interchangeable parts that can function in conjunction with cells natural systems. Natural systems are comprised of complex networks that are regulated at different levels, including transcriptional and translational regulation. However, most synthetic circuits are governed at the transcriptional level; very few utilize translational regulation[4]. Utilization of native translational regulation in synthetic circuits holds potential that has yet to be explored. The translational level control is beneficial as it promotes rapid output that allows fine-tuning of the system.

We recently developed a new class of synthetic circuits that can be controlled at both the transcriptional and translational level mimicking native biological networks. These circuits include an AND logic gate, which requires two inputs, one at the transcriptional level and one at the translational level, to activate protein production of the output (typically a fluorescent protein for quantitative data acquisition). Our circuit exploits cells' innate phenomenon called ribosome stalling, which also regulates translation rate of many natural transcripts[6-7]. It is well established that when a ribosome encounters clusters of rare codons it stalls, but the ribosome can be "rescued" by tmRNA. The tmRNA allows the ribosome to abort translation of the current peptide chain via the simultaneous completion of the polypeptide with a fast-degradation tag and cleavage/degradation of the mRNA molecule at the site of stalling[7-9]. Ribosome rescue is a fundamental process that allows cells to recover unproductive ribosomes when the cell is under stress and lacks a sufficiently large number of charged (i.e. bound to an amino acid) tRNA molecules for translation.

Our recently developed AND gate leverages sequential rare codons for differential production of fluorescent proteins at the translational level. This construct is under an inducible promoter with transcriptional control and translational level control determined by the availability of cognate rare tRNA in the cell, which is expressed under a different inducible promoter. The system's feasibility is tested in stationary phase in population study that shows significant fold difference in fluorescence level when AND logic gate is functional. Further analysis of this system will be performed in single cell study using microfluidics to gather quantitative information.

Our goal is to develop scalable, robust synthetic circuit, which can act as interchangeable parts in a wide range of organisms. We are planning to modify the current system to increase its robustness and move towards building additional logic gates using available parts in order to build a complex network in biological native system.

II. References

[1] Gardner, T. S.; Cantor, C. R.; Collins, J. J., Construction of a genetic toggle switch in Escherichia coli. Nature 2000, 403 (6767), 339-42.

[2] Tyson, J. J.; Chen, K. C.; Novak, B., Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Curr Opin Cell Biol 2003, 15 (2), 221-31.

[3] Guet, C. C.; Elowitz, M. B.; Hsing, W.; Leibler, S., Combinatorial synthesis of genetic networks. Science 2002, 296 (5572), 1466-70.

[4] Gao, X. J.; Chong, L. S.; Kim, M. S.; Elowitz, M. B., Programmable protein circuits in living cells. Science 2018, 361 (6408), 1252-1258.

[5] Sherlock, M. E.; Sudarsan, N.; Breaker, R. R., Riboswitches for the alarmone ppGpp expand the collection of RNA-based signaling systems. Proceedings of the National Academy of Sciences 2018, 115 (23), 6052-6057.

[6] Konigsberg, W.; Godson, G. N., Evidence for use of rare codons in the dnaG gene and other regulatory genes of Escherichia coli. Proceedings of the National Academy of Sciences 1983, 80 (3), 687-691.

[7] Roche, E. D.; Sauer, R. T., SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity. EMBO J 1999, 18 (16), 4579-89.

[8] Mitarai, N.; Sneppen, K.; Pedersen, S., Ribosome collisions and translation efficiency: optimization by codon usage and mRNA destabilization. Journal of molecular biology 2008, 382 (1), 236-45.

[9] Janssen, B. D.; Hayes, C. S., The tmRNA ribosome-rescue system. Advances in protein chemistry and structural biology 2012, 86, 151-91.

1. Department of Biology and Microbiology, South Dakota State University, Brookings, South Dakota 57007, United States. E-mail: prajakta.jadhav@sdstate.edu, nicholas.butzin@sdstate.edu.

Modeling epigenetic feedback regulation during Epithelial-Mesenchymal Transition (EMT)

Wen Jia[1], Mohit Kumar Jolly[2], Abhijeet Deshmukh[3], Sendurai A. Mani[3] and Herbert Levine[1,4]

***Short Abstract* — Epithelial-mesenchymal transition (EMT) plays an important role in cancer metastasis and drug resistance, and involves epigenetic remodeling However, how epigenetic changes affecting the dynamical traits such as plasticity or memory are not fully understood. Here, we analyze the effects of epigenetic feedback on EMT through integrating this feedback on various aspects of the miR-200/ZEB loop – a core circuit regulating EMT. Epigenetic feedback on self-activation of ZEB has minor effects in population distribution and transition times, but epigenetic feedback on the inhibition of miR-200 by ZEB can largely stabilize the mesenchymal state, thus making the process irreversible. Follow-up preliminary experiments show that when EMT is induced in epithelial cells , a certain percentage of cells can stay in mesenchymal state after the inducing signal is removed. This percentage depends on the extent of induction of EMT, thus well recapitulating our model-based predictions.**

***Keywords* — EMT, epigenetic feedback**

I. BACKGROUND

THE epithelial-mesenchymal transition(EMT) and its reverse transition(MET) are closely related to cancer metastasis - the spreading of cancer cells from original tumor to a different part of the body and forming a secondary tumor.

Epigenetic factors can affect EMT by modulating gene expression through altering chromatin structure to create active or inactive transcription regions [1, 2]. A recent study indicates that inhibition of specific histone acetylation and demethylation pathways can suppress the induction as well as the maintenance of SNAIL-driven EMT [3]. However, the dynamical features of such epigenetic modifications remain unclear. Here, to better understand the effect of epigenetics on EMT, we introduced an epigenetic feedback term in a core EMT circuit that consists of two mutually inhibiting loops: miR-200/ZEB and miR-34/SNAIL [4].

II. RESULTS

The core EMT circuit miR-200/ZEB/miR-34/SNAIL allows for the co-existence of three stable states – epithelial, mesenchymal and hybrid E/M [4]. Through stochastic simulations, we quantified the population distribution in this tristable system for the abovementioned core circuit and its reduced version miR-200/ZEB/SNAIL, and analyzed the effects of epigenetic feedback on altering this distribution and transition rates among various phenotypes.

A. Epigenetic feedback on ZEB's self-activation

The population distributions almost remain unchanged when epigenetic feedback is on the self-activation of ZEB, even if the feedback is very strong.

B. Epigenetic feedback on ZEB's inhibition on miR-200

From both bifurcation and population distribution results, we can see that the mesenchymal state now becomes greatly stabilized, and more epithelial cells are able to reach and maintain the mesenchymal state. If the epigenetic feedback on the inhibition on miR-200 by ZEB is strong enough, more than 50% of cells that maintained an epithelial cells in case of no feedback case can now transition to a mesenchymal state.

C. Experimental results

Our model predicts that with epigenetic feedback on the inhibition on miR-200 by ZEB, some cells can maintain their mesenchymal state after the external stimulus is removed. This irreversibility is seen in preliminary experimental data of cells treated with EMT-inducing signals for varying duration.

III. CONCLUSION

EMT system is not sensitive to the epigenetic feedback on self-activation of ZEB. However, epigenetic feedback on inhibition on miR-200 of ZEB can stabilize mesenchymal state and contribute to the irreversibility of EMT, thus potentially restricting formation of metastases. Our results demonstrated that targeted epigenetic regulation in EMT may affect the ability of cells to complete their metastatic cascade.

REFERENCES

[1] Ernst, J., Kheradpour, P., Mikkelsen, T. S., Shoresh, N., Ward, L. D., Epstein, C. B., ... & Ku, M. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*, *473*(7345), 43.

[2] Marstrand, T. T., & Storey, J. D. (2014). Identifying and mapping cell-type-specific chromatin programming of gene expression. *Proceedings of the National Academy of Sciences*, *111*(6), E645-E654.

[3] Javaid, S., Zhang, J., Anderssen, E., Black, J. C., Wittner, B. S., Tajima, K., ... & Maheswaran, S. (2013). Dynamic chromatin modification sustains epithelial-mesenchymal transition following inducible expression of Snail-1. *Cell reports*, *5*(6), 1679-1689.

[4] Lu, M., Jolly, M. K., Levine, H., Onuchic, J. N., & Ben-Jacob, E. (2013). MicroRNA-based regulation of epithelial–hybrid–mesenchymal fate determination. *Proceedings of the National Academy of Sciences*, 201318192.

[1]Center for Theoretical Biological Physics, Rice University, Houston, TX, USA. E-mail: wj7@rice.edu

[2]Centre for BioSystems Science and Engineering, Indian Institute of Science, Bangalore, India.

[3]Department of Translational Molecular Pathology, University of Texas MD Anderson Cancer Center, Houston, TX, USA.

[4]Department of Physics, Northeastern University, Boston, MA, USA.

Mechanics before Chemistry: Tensile Stress Induced Cytoskeletal Reorganization

Xiaona Li[1], Qin Ni[2], Xiuxiu He[1], Soon Mi Lim[3], Garegin A. Papoian[2], Andreea Trache[3,4] and Yi Jiang[1]

***Short Abstract* — Cellular remodeling in response to mechanical stimuli is critical for understanding mechano-signal transduction. We hypothesize that external stress induced subcellular adaption is accomplished through dynamical cytoskeletal reorganization. To study the interactions between subcellular structures involved in transducing mechanical signals, we combined experimental and modeling approaches to measure real-time structural and mechanical adaption of the actin cytoskeletal network. In vitro, we imaged the actin cytoskeleton as tensile stress was applied to live vascular smooth muscle cells (VSMC) using an ECM-functionalized atomic force microscope probe. In silico, we modeled the mechanochemical coupling of the actin cytoskeleton network. Both experimental and modeling results agree that under tensile stress, mechanical structural adaptation occurs before chemical adaptation: actin filaments align first, then actin polymerization takes place to further restructure the cytoskeleton.**

***Keywords* — Cytoskeletal Network, Tensile Stress, Filament Alignment, Actin Polymerization**

I. Introduction

CELLS interact with a complex microenvironment. Among all the microenvironmental stimuli, mechanical stress [1, 2] is important in many biological and physiological processes. Vascular smooth muscle cells (VSMC) are subjected to the cyclic stretch of pulsatile blood pressure that deforms the extracellular matrix and induces axial and circumferential wall stresses [3]. However, how VSMC responds to the mechanism of axial stress in the vessel wall, which can be considered as tensile stress applied to cell, is not well-understood [4]. We combine experimental and modeling approaches to investigate the effects of tensile stress on the dynamic remodeling of the cytoskeleton network.

II. Methods

The tensile stress was applied to live VSMC using an atomic force microscope probe functionalized with extracellular matrix proteins. Mechanical stimulation of the cell at low (~0.5 nN) and high (~1 nN) magnitude forces was applied every 3–5 min for 20–25 min each and the actin cytoskeleton was imaged by spinning-disk confocal microscopy after each force application [3].

A computational model for mechanochemical dynamics of active network (MEDYAN) [5] was used to simulate the actin network with an external pulling force. The model considers actin fibers as semi-flexible polymers embedded in a solution of actin monomers, alpha-actinin cross-linking proteins, and non-muscle myosin II (NMII) motors. A system of reaction-diffusion equations describes the spatiotemporal dynamics of actin polymerization and actomyosin network formation. We initialized 300 seed filaments randomly generated in a 3×3×1 μm^3 simulation volume, where these filaments are allowed to polymerize and depolymerize on both plus ends and pointed ends. Additionally, we created a 250 nm in radius semi-sphere to mimic AFM tip. The AFM tip applies an external pulling force. We varied the strength of force, and measured the resulting fiber alignment, polymerization, and uniformity, and compare with experimental measurements.

III. Results

Both experimental and simulation results show that tensile stress has significant effect on the dynamics of the cytoskeleton network: as the tensile stress increases, the fibers rearrange to increase alignment along the direction of the external stress, before fiber polymerization takes place. This result suggests mechanical structural adaptation operates at a shorter timescale than biochemical processes, which can have important implications to mechano-signal transduction.

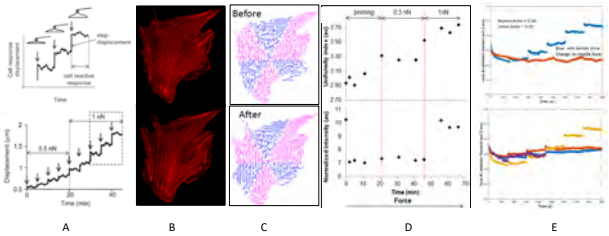


Figure 1. A. Application of external force using AFM on VSMC. B. Representative images of VSMC expressing actin-mRFP before (top) and after external force. C. Segmented acting fiber orientations. D. Uniformity and intensity actin as a function of time. F. Simulated fiber alignment as a function of time with and without external force.

References

[1] Lu, D., & Kassab, G. S. (2011). *J. Royal Soc. Interface*, 8(63), 1379-1385.
[2] Li, C., & Xu, Q. (2007). *Cell Signal*, 19(5), 881-891.
[3] Lim, S. M. et al. (2012). *Integr. Biol.*, 4(6), 615-627.
[4] Patel, D. J. et al. (1966).. *Circ. Res.*, 19(6), 1011-1021.
[5] Popov, K. et al. (2016). *PLoS Comput.l Bio.*, 12(4), e1004877.

[1]Department of Mathematics and Statistics, Georgia State University, Atlanta, GA, USA. E-mail: yjiang12@gsu.edu
[2]Department of Chemistry & Biochemistry, University of Maryland, College Park, MD, USA.
[3]Department of Medical Physiology, Texas A&M Health Science Center, College Station, TX, USA. E-mail: trache@tamu.edu
[4]Department of Biomedical Engineering, Texas A&M University, College Station, TX, USA.

Dynamics of Viral Evolution

Greyson Lewis[1], Wallace Marshall[2], and Barbara A Jones[3]

Short Abstract **-- We study a model of viral evolution; viruses have a barrier to cell entry, via match to cell "key," followed by cell immune response, and probability to reproduce and mutate. These mutated viruses then attack other cells in the model. Our steady state results feature a phase transition separating viruses with different survival strategies. Here we describe our computational studies of the behavior of this model as a dynamical system, and the nonequilibrium evolution of the quasispecies distribution including metastable states and other unexpected features. We have found evidence of the competition between strategies in all dynamical properties.**

Keywords **— Viruses, evolution, mutation, dynamics, phase transition**

Viruses infect cells of living organisms across all types of life.[1] To replicate, they must enter a host cell and use its DNA. We study an idealized model of viral mutation and evolution in human cells, which displays a richness of behavior both in steady state [2] and in its dynamics. In this model, viruses first have a barrier to entry. Viruses are assigned a match number, according to how well they match the genetic receptor target "key" on the cell. The probability of successful cell barrier crossing has an activated Arrhenius form depending on the viral match number and temperature.

If a virus enters the cell, there is an immune response, If a virus has been successful in the past in entering the cell, the immunity remembers it and has a stronger immune response, modulated by amplitude A. Finally the virus attempts to reproduce and exit the cell. We take the probability of replication, mutation, and cell death to be also an activated form, including a probability to remain in the cell. We have one mutation per viral offspring, changing m. Thus we have an inherent tension in our model: if a virus maximizes its chance for entry into the cell, it is more likely to be killed by the immunity. If it tries to avoid the immunity, it has less chance of reproduction.

The host of viruses from the cells then becomes the new environment, and goes on to try to infect the cells again. This cycle of infection, immune response and reproduction/mutation repeats. By this process the distribution and number of viruses in the environment changes, as does the population and distribution of viruses in the cells. Depending on the initial parameters T (temperature) and A (immune response), the viruses can take many different strategic paths to survival. After many iterations, the system goes to one and only one steady state determined by the choice of T and A.

Summary of Steady State Results: The steady state solutions of our model [2] show a phase transition evident in the order parameter, taken to be the number of matches. This phase transition separates strategies of small or large average match number. These can be viewed mimicking two strategies of real viruses. Some, like the cold or flu, come in strong and then are cut down by the immunity and eventually (mostly) cleared from the body. Others, like chicken pox, stay in the body and can emerge 60 years later as shingles as immunity decreases with age.

Dynamics: Now we turn to the dynamics, and investigate the potentially interesting biological and mathematical behavior, either independent of or reflecting the steady state. The dynamical process of the many iterations as the viruses mutate on their way to steady state may in fact be the dominant behavior seen in a natural human infection cycle, especially if the time to steady state is long.

We calculate a range of properties, such as fitness, average match number, and viral population. In each case we find that there are evident different strategies as T and A are varied, and that these strategies divide along the same line that separates the phases in steady state. We see evidence of cooperation, competition, bimodal distributions, and intra-quasispecies population shifts, a complex interplay of the viruses as they evolve toward maximum fitness.

In addition, if the initial values of T and A are near the eventual phase transition line, the number of iterations to steady state can go into the hundreds of thousands or more. We see metastable states as the viruses try to find optimal properties in a complex energy/entropy landscape. We show how this behavior is indicative of glassy behavior.

Finally, our system sizes involve hundreds of thousands of cells and viruses, and I will describe some of the big data challenges arising from these calculations.

REFERENCES

[1] Koonin EV, Senkevich TG, Dolja VV. (2006) The ancient virus world and evolution of cells. Biol Direct. 1:1.
[2] Jones BA, Lessler J, Bianco S, Kaufman JH (2015) Statistical Mechanics and Thermodynamics of Viral Evolution. PLoS ONE 10(9): e0137482.e

[1] Biochemistry and Biophysics, School of Medicine, UCSF.
E-mail: greyson.lewis@gmail.com
[2]Biochemistry and Biophysics, School of Medicine, UCSF
E-mail: Wallace.Marshall@ucsf.edu
[3]IBM Research Almaden, San Jose, CA. E-mail: bajones@us.ibm.com

Computational Design of Synthetic Microbial Communities

Behzad D. Karkaria[1], Alex J. Fedorec[2] and Chris P. Barnes[3]

***Abstract*— In naturally occurring microbial systems, species rarely exist in isolation. Building controlled communities for synthetic biology has advantages for genetic insulation, and the potential to create distributed systems, reducing burden and enabling greater complexity. The existence and maintenance of consortia is dependent upon the interactions between community members. Here, we demonstrate a modelling framework for simulating microbial consortia existing in a chemostat environment, interacting via quorum sensing and microcin expression. Using this framework, we are able to predict system configurations that give the highest probability for stable steady state communities, and infer the design principles underlying complex community design.**

I. Extended Abstract

Traditionally in synthetic biology and biotechnology, a microbe is engineered as a monoculture to perform a heterologous function. Each heterologous process comes at a metabolic cost to the host. The metabolic capacity of a single host therefore, imposes constraints upon the complexity of a monoculture system [1]. In natural environments we observe mixed-species consortia, exhibiting competitive advantages over monocultures in productivity, resource efficiency, metabolic complexity and resistance to invasion [2], [3], [4]. A consortium also brings the benefits of compartmentalised sub-processes that are distributed between sub-populations of the community. This enables higher modularity, insulation from inter-circuit interference, and relief from the the restriction of host metabolic burden [5]. The ability to predictably and reproducibly establish microbial communities would allow us to apply these advantages to synthetic biology and biotechnology.

Interactions between subpopulations of a microbial consortium are essential for its maintenance and robustness. We can engineer communication between subpopulations through quorum sensing systems [6]. We can also engineer feedback systems via microcins, causing ammensal interactions between subpopulations. For a given set of quorum systems and microcins, there are different combinations in which these can be assembled into a system. Here we present a computational work flow that allows us to specify a set of parts that defines the model space, and automatically generate mathematical models for all possible systems. We simulate bacterial growth, quorum molecule expression and quorum induced microcin expression for each of these systems, classifying the behaviours found. This allows us to perform model selection to identify optimal topologies, infer the design principles and identify experimental parameters necessary to establish stable steady state communities.

[1]Department of Cell and Developmental Biology, University College London. Email: behzad.karkaria.16@ucl.ad.uk

[2]Department of Cell and Developmental Biology, University College London. Email: alexander.fedorec.13@ucl.ac.uk

[3]UCL Genetics Institute, UCL Genetics Institute, University College London. E-mail: christopher.barnes@ucl.ac.uk

II. Conclusion

In summary, we have developed modelling framework that allows us to assess model spaces consisting of multiple strains expression quorum sensing systems and quorum controlled microcins. We demonstrate the use of this framework for making model informed decisions in the design of engineered microbial communities.

References

[1] A. Y. Weiße, D. A. Oyarzún, V. Danos, and P. S. Swain, "Mechanistic links between cellular trade-offs, gene expression, and growth." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 9, pp. E1038–1047, 2015. [Online]. Available: http://www.pnas.org/content/112/9/E1038.abstract

[2] V. H. Smith, "Microbial diversity productivity relationships in aquatic ecosystems," *FEMS Microbiology Ecology*, vol. 62, no. 2, pp. 181–186, nov 2007. [Online]. Available: https://academic.oup.com/femsec/article-lookup/doi/10.1111/j.1574-6941.2007.00381.x

[3] D. Beyter, P.-Z. Tang, S. Becker, T. Hoang, D. Bilgin, Y. W. Lim, T. C. Peterson, S. Mayfield, F. Haerizadeh, J. B. Shurin, V. Bafna, and R. McBride, "Diversity, Productivity, and Stability of an Industrial Microbial Ecosystem." *Applied and environmental microbiology*, vol. 82, no. 8, pp. 2494–2505, apr 2016. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/26896141http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4959475

[4] J. D. van Elsas, M. Chiurazzi, C. A. Mallon, D. Elhottova, V. Kristufek, and J. F. Salles, "Microbial diversity determines the invasion of soil by a bacterial pathogen." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 4, pp. 1159–64, jan 2012. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/22232669http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3268289

[5] K. Brenner, L. You, and F. H. Arnold, "Engineering microbial consortia: a new frontier in synthetic biology," *Trends in Biotechnology*, vol. 26, no. 9, pp. 483–489, 2008.

[6] N. Kylilis, Z. A. Tuza, G.-B. Stan, and K. M. Polizzi, "Tools for engineering coordinated system behaviour in synthetic microbial consortia," *Nature Communications*, vol. 9, no. 1, p. 2677, dec 2018. [Online]. Available: http://www.nature.com/articles/s41467-018-05046-2

Translation bottlenecks underlie diverse drug interactions between translation inhibitors

Bor Kavčič[1,*], Tobias Bollenbach[2] and Gašper Tkačik[1]

***Abstract*— Antibiotic combinations are important for treating infections and as tools in basic research. Drug interactions – determined by the potency of the drug combination – can reveal important relations in cell physiology. We experimentally show that drug interactions between translation inhibitors largely stem from the interplay of various translation bottlenecks. We show that suppressive drug interactions are caused by ribosomal traffic jams, which are alleviated by lowering the initiation rate. We explain this effect using a generalized TASEP model of translation that takes bacterial growth laws into account. This work reveals novel mechanisms of drug interactions and elucidates the coordination of translation.**

***Index Terms*— Antibiotics, growth laws, TASEP, translation, translation factors, antibiotic interactions.**

I. Introduction

Although antibiotics have been used in clinical settings for almost a century, bacterial responses to antibiotics are not well understood. It is important to understand these response mechanisms as resistance to the current arsenal of antibiotics is becoming a significant problem. A promising way to combat antibiotic resistance is combination therapy in which multiple antibiotics are applied simultaneously. Responses to drug combinations are generally difficult to predict and include synergism (inhibition is stronger than predicted), antagonism (inhibition is weaker) and suppression (one of the drugs loses potency). Apart from the clinical relevance, antibiotic combinations are important for understanding the interplay between different cellular systems – similar to studies of epistasis among double gene knockouts.

Growth of bacteria requires a well-orchestrated translation machinery which is modulated internally by translation factors and perturbed by certain antibiotics (translation inhibitors). Coupling of the growth rate to ribosome synthesis results in ribosomal growth laws: a phenomenological, yet precise feedback model for how growth determines the total number of ribosomes [1]. Mathematical models have been used for successful quantitative predictions of the growth rate in the presence of antibiotics as a function of their concentration [2]. However, a similar understanding of the combined action of antibiotics is still lacking.

II. Results

We began by measuring the drug interaction network for a set of translation inhibitors, which revealed a diversity of interactions. We developed a mathematical model of bacterial growth in the presence of combinations of antibiotics targeting the ribosome. Our theoretical description captures key physical processes (transport and binding of antibiotics to the ribosome) as well as the growth laws, yet depends on only a single measurable parameter per antibiotic. The theoretical model explained some of the interactions, but was unable to explain the occurrence of a severe antagonism and suppression. To further examine the origin of these unexplained drug interactions, we mimicked antibiotic effects on translation genetically by externally controlling the expression of one or several key translation factors, which revealed how antibiotic action depends on the translation bottlenecks. Firstly, we find that antibiotic responses to bottlenecks cluster into functional groups, allowing us to corroborate the mode of action of a recently identified small molecule that inhibits initiation [3]. Secondly, we show how to predict pairwise drug interactions from measured responses of antibiotics to translation bottlenecks, using a mapping that we call *interventions of equivalent effect*. Thirdly, we show that suppressive interactions between antibiotics targeting initiation and translocation arise from traffic jams of the ribosomes in the translation cycle. Traffic jams can arise when the ribosome transition rates are modified, leading to a decrease in translation efficiency, an effect we theoretically account for in a stochastic model of translation based on the totally asymmetric exclusion process (TASEP).

III. Conclusion

We developed a theoretical and experimental framework for the analysis of antibiotic interaction mechanisms. Our results suggest that translation bottlenecks in the context of growth laws underlie the diverse drug interactions between translation inhibitors. This work connects single-antibiotic studies to mechanism-independent models of combined multidrug responses by providing a novel methodology for prediction of pairwise drug interactions [4].

References

[1] M. Scott *et al.*, "Interdependence of cell growth and gene expression: Origins and consequences," *Science*, vol. 330, 2010.

[2] P. Greulich *et al.*, "Growth-dependent bacterial susceptibility to ribosome-targeting antibiotics," *Molecular systems biology*, vol. 11, 2015.

[3] J. Stokes *et al.*, "Discovery of a small molecule that inhibits bacterial ribosome biogenesis," *eLife*, vol. 3:e03574, 2014.

[4] K. Wood *et al.*, "Mechanism-independent method for predicting response to multidrug combinations in bacteria," *PNAS*, vol. 109, 2012.

[1]Institute of Science and Technology Austria, Klosterneuburg, Austria,*Email: bor.kavcic@ist.ac.at

[2]University of Cologne, Cologne, Germany

Tumor cell phenotype and heterogeneity differences in IDH1 mutant vs wild-type gliomas

Michael E. Berens[1], Anup Sood[3], Jill S. Barnholtz-Sloan[2], John F. Graf[3], Sanghee Cho[3], Seungchan Kim[4], Jeff Kiefer[1], Sara A. Byron[1], Rebecca Halperin[1], Sara Nasser[1], Jonathan Adkins[1], Lori Cuyugan[1], Karen Devine[2], Quinn Ostrom[2], Marta Couce[2], Leo Wolansky[2], Winnie Liang[1], Mirabela Rusu[3], Maria I Zavodszky[3], Fiona Ginty[3]

***Short Abstract* — Glioma is a highly heterogeneous malignancy, whose diverse cellular composition and cellular interactions have not been well characterized. To glean new clinically- and biologically-relevant insights into IDH1 mutant (mt) vs wildtype (wt) disease, we integrated multiplexed immunofluorescence single cell data from over 43 protein markers of cell types, cancer hallmarks, dysregulated pathways, cell spatial metrics, cancer genomic sequencing and magnetic resonance imaging metrics. Correlations between imaging, cell clusters and molecular features at the hallmark level were observed. Delineating biological events underlying glial tumors in multiscale and spatial context may facilitate tailored treatment approaches as well as reveal new therapeutic targets.**

I. BACKGROUNDS

Gliomas represent the most common type of malignant brain tumor, with significant morbidity and mortality; however, our current understanding of its diverse cellular heterogeneity and interactions is significantly limited. Intratumoral heterogeneity is a characteristic of glioma and glioblastoma (GBM) (1), occurring both temporally (2) and spatially (3). IDH1 mutation status has been shown to determine G-CIMP, a DNA methylation-based phenotype, which is predictive of increased survival. Medical image and radiomics metrics have been used to identify both cellular heterogeneity and differentiation of IDH1mt vs wt (4). Understanding malignant progression in IDH1mt and wt patients at multiple scales and in a spatial context is pivotal to delineating biological events underlying glial tumors.

II. RESULTS

In order to discern cellular composition and spatial differences in relation to IDH1 mutation status, we conducted a multiscale integrative analysis of multiplexed immuno-fluorescence and single cell spatial analysis of fixed glioma tissue, genomic tumor sequencing, MR imaging data of the whole tumor and sub-regions, and patient outcomes. Separate multiscale datasets were assembled from treatment-naïve cases of grade 2-4 astrocytoma/oligodendroglioma (n=20) and from recurrent grade 4 astrocytoma (GBM) (n=16).

We show lower cell-level hallmark-associated protein expression in IDH1mt vs wt cases. Further, IDH1mt gliomas, irrespective of grade, showed greater spatial heterogeneity of biomarkers associated with angiogenesis (VEGR2, CD31, SMA, S100A4) and invasion (n-cadherin, cofilin, collagen IV, GFAP and vimentin). Molecular heterogeneity was significantly lower in angiogenesis and invasion markers in IDH1mt tumors. Similarly, cell classes derived from deconvolution of bulk gene expression data showed that cell classes with high expression of most hallmark genes, particularly those belonging to enabling replicative immortality, evading growth suppressors and inducing angiogenesis, were significantly under represented (<10%) in the IDHmt tumors. IDH mutation was co-expressed with ATRX mutations and was mutually exclusive of EGFR and PTEN mutations consistent with known tumor biology. Longer overall survival following diagnosis for IDH1mt glioma patient may reflect generalized altered cellular, molecular and spatial heterogeneity, which is also reflected in the MR images as reduced enhancement core volume and higher contrast uptake in the peritumoral edema region.

III. CONCLUSION

In this study, we evaluated the correlation between imaging and molecular features at the hallmark level, via an integrative workflow for multiscale, multiparametric data. The similarity of cell clusters between primary glioma and recurrent GBM is surprising and instructive of enduring multiscale pathobiology.

REFERENCES

1. Aum DJ, Kim DH, Beaumont TL, Leuthardt EC, Dunn GP, Kim AH. Molecular and cellular heterogeneity: the hallmark of glioblastoma. Neurosurg Focus. 2014;37(6):E11. doi: 10.3171/2014.9.FOCUS14521.
2. Sottoriva A, Spiteri I, Piccirillo SG, Touloumis A, Collins VP, Marioni JC, et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proc Natl Acad Sci U S A. 2013;110(10):4009-14. doi: 10.1073/pnas.1219747110.
3. Kumar A, Boyle EA, Tokita M, Mikheev AM, Sanger MC, Girard E, et al. Deep sequencing of multiple regions of glial tumors reveals spatial heterogeneity for mutations in clinically relevant genes. Genome Biol. 2014;15(12):530. doi: 10.1186/s13059-014-0530-z.
4. Ryu YJ, Choi SH, Park SJ, Yun TJ, Kim JH, Sohn CH. Glioma: application of whole-tumor texture analysis of diffusion-weighted imaging for the evaluation of tumor heterogeneity. PLoS One. 2014;9(9):e108335. doi: 10.1371/journal.pone.0108335.

Acknowledgements: This work was partly funded by NIH grant 1U01CA168397 (MB, SK) and Texas A&M System Chancellor's Research Initiative for Center for Computational Systems Biology (SK).

[1]Translational Genomics Research Institute, Phoenix, AZ, USA.
[2]Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, Ohio, USA.
[3]GE Research Center, Niskayuna, NY, USA.
[4]Prairie View A&M University, Prairie View, TX, USA

Modeling Signal Cross-talk in Type I Interferon

Duncan Kirby[1], Grégoire Altan-Bonnet[2] and Anton Zilman[1,3]

***Abstract*— Cross-wired signaling pathways are common in cell signaling, making it unclear how cells discriminate between different ligands. This issue features prominently in Type I Interferon (IFN) signaling, a process crucial to the innate immune response. Using theoretical modeling and quantitative experiments, we systematically investigate factors responsible for specificity of Type I IFN signaling. We show that specificity arises from the interplay of receptor binding affinity, the numbers of receptor subunits and their in-membrane mobility, amplified by transcriptional negative feedback. We also formulate a general theoretical framework which shows how multiple ligands can be accurately sensed in the presence of cross-talk.**

***Index Terms*— modeling, cytokines, systems immunology, immune system**

I. Background

Biochemical signaling networks enable cells to respond to changes in their environment, whether for survival or for tissue function. Traditionally, these signaling networks were studied as direct input-output relationships. This view is now undermined by frequent observation of signal cross-talk, a phenomenon wherein several signals are transmitted through interacting signaling pathways. This can create complicated signaling logic rules. Studying these networks experimentally is difficult since measuring cellular response to isolated signals misses complex signaling patterns due to cross-talk, and experiments with multiple signals make causal relations hard to infer. Mathematical models have recently shown how nonlinear dynamical modeling and statistical inference can help to overcome these challenges [1], [2].

A prime example of cross-talk is a family of innate immune signaling molecules called Type I Interferons, which includes Interferon-α (IFNα) and Interferon-β (IFNβ). IFNα classically induces an antiviral response while IFNβ induces an antiproliferative response. It is not well understood how a cell achieves these different responses since these Interferons share a cell surface receptor (a prototypical case of cross-talk). Many factors affecting signal specificity of Type I IFN have been identified in the literature including binding affinities [3], dynamic receptor expression [4], receptor internalization [5], and protein-mediated negative feedback on several timescales [6]. How these various factors interplay to affect signal specificity has not been systematically investigated until now.

[1]Department of Physics, Faculty of Arts and Sciences, University of Toronto, Toronto, ON M5S 1A7, Canada

[2]Immunodynamics Group, Cancer and Inflammation Program, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20814, USA

[3]Institute for Biomaterials and Bio-engineering, Faculty of Engineering, University of Toronto, Toronto, ON M5S 1A7, Canada

II. Results

We have developed a mathematical model of Type I Interferon signaling which captures all of the essential features of this signaling network. Our model has been validated by quantitative flow cytometry measurements of cellular response to IFNα and IFNβ. We find that IFN-specific negative feedback such as receptor internalization is necessary to explain observed differences between Interferons. We have also used a combination of stochastic modeling and maximum likelihood estimation to formulate bounds on the accuracy and specificity of signaling in the presence of cross-talk. This theoretical work shows that receptor pleiotropy overcomes fundamental limitations in sensing multiple ligands in the presence of cross-talk. The resulting bounds on signaling accuracy and specificity offer a way to benchmark the performance of Interferon signaling according to our dynamical model.

III. Significance

A quantitative model of cross-talk signaling is necessary to discover response rules for mixtures of signals, essential for studying such signaling processes. Furthermore, these models are invaluable to drug design research where cross-talk can alter pharmacological effects. This research will open the door to new immune therapies for disease.

References

[1] S. Fathi *et al.*, "Absolute ligand discrimination by dimeric signaling receptors," *Biophys J*, vol. 111, pp. 917–920, 2016.

[2] M. Carballo-Pacheco *et al.*, "Receptor crosstalk improves concentration sensing of multiple ligands," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 99, no. 2, pp. 1–14, 2019.

[3] G. Schreiber, "The molecular basis for differential type i interferon signaling," *Journal of Biological Chemistry*, vol. 292, no. 18, pp. 7285–7294, 2017.

[4] D. Levin, D. Harari, and G. Schreiber, "Stochastic receptor expression determines cell fate upon interferon treatment," *Molecular and Cellular Biology*, vol. 31, pp. 3252–66, 2011.

[5] Z. Marijanovic *et al.*, "Comparable potency of IFNα2 and IFNβ on immediate JAK/STAT activation but differential down-regulation of IFNAR2," *Biochemical Journal*, vol. 407, no. 1, pp. 141–151, 2007.

[6] S. Wilmes *et al.*, "Receptor dimerization dynamics as a regulatory valve for plasticity of type I interferon signaling," *Journal of Cell Biology*, vol. 209, no. 4, pp. 579–593, 2015.

History-dependent Maintenance of Drug Resistant Phenotypes against Resistant Gene Deletion

Yuta Koganezawa[1], Moritoshi Sato[1,2,3] and Yuichi Wakamoto[1,2,3]

***Short Abstract*— How rapidly cells transit to a new terminal phenotype after the introduction of some genetic modification and how uniform the phenotypic transition processes are at the single-cell level are both important questions in genetics but have not been explored in details. Here, we introduce an approach that combines microfluidic device and optogenetic gene recombination system and show that 40% of *Escherichia coli* cells that lost chloramphenicol (Cp) resistant gene in Cp-exposed environment nevertheless continued stable growth and division. The maintenance of the resistance against gene deletion occurred in Cp-exposed environment, which suggests the history-dependent nature of the genotype-phenotype mapping.**

***Keywords*— genotype-phenotype mapping, optogenetical gene recombination, single-cell observation, drug resistance**

I. Introduction

Understanding how tightly genotypes constrain phenotypes of an organism in a given environment is a fundamental problem in genetics. Genotype-phenotype mapping is usually investigated by comparing the terminal and stable phenotypes between the cells with a reference genotype and those with some genetic changes. In recent years, the availability of knock-out library strains, high-throughput imaging techniques, and automated image analysis enabled us to explore the linkage between genotypes and phenotypes extensively[1], [2]. However, detailed studies on the phenotypic transition towards the terminal phenotypes that occur after the introduction of genetic modification are scarce despite some interests in the 1940s to 1950s[3]. Furthermore, it is even less examined how uniform or heterogeneous the transition processes are at the single-cell level.

To address this, we monitored the responses of individual cells to genetic modification combining Mather Machine, a microfluidic device for single-cell observation[4], and Magnet-Cre, a genetic recombination system inducible by blue light[5]. Magnet-Cre system consists of split Cre recombinase fragments tagged with p- and n-Magnet monomers engineered from a fungal photoreceptor. Blue light illumination induces heterodimerization of p- and n-Magnet, which recovers Cre recombination activity. In this study, we provoked pre-designed deletion of fluorescently-tagged and chromosomally-encoded chloramphenicol(Cp) resistant gene, *mcherry-chloramphenicol acetyltransfease* (*mcherry-cat*), in *Escherichia coli* at arbitral timings by blue-light activation of Magnet-Cre system.

[1]Department of Arts and Sciences, The University of Tokyo
[2]Research Center for Complex Systems Biology, The University of Tokyo
[3]Universal Biology Institute, The University of Tokyo

II. Results

We first checked the efficiency and toxicity of genetic recombination by Magnet-Cre in a Cp-free environment in Mother Machine, finding that 30-minute blue-light exposure caused intended gene deletion in 20% of the cells and that there was no significant difference in the growth rates between the cells before and after the gene deletion.

Next, we induced the deletion of *mcherry-cat* under the conditions of continuous exposure of Cp at a lethal dosage. The result showed that 40% of the cells which lost *mcherry-cat* gene continued growth and division stably, i.e. preserved the resistant phenotype, for at least 30 generations without the resistant gene. Changing the timings of starting Cp exposure revealed that all the cells stopped growing and dividing when Cp exposure started 10 hours after the gene deletion. These results indicate that maintenance of the resistant phenotype occurred depending on the environmental histories and the timing of gene deletion.

III. Conclusion

We observed the phenotypic transition processes at the single-cell level that occurred in response to the optogenetic deletion of the drug resistance gene. *E. coli* cells maintained the resistance against the removal of the resistance gene when Cp was present in the environment. The results demonstrate the history-dependent genotype-phenotype mapping and the unexpectedly long-term memory of vital phenotypes.

References

[1] S. F. Levy, M. L. Siegal. Network hubs buffer environmental variation in Saccharomyces cerevisiae. *PLoS Biology*, 6(11):e264, 2008.
[2] M. Campos *et al*. Genomewide phenotypic analysis of growth, cell morphogenesis, and cell cycle events in Escherichia coli. *Molecular Systems Biology*, 14(6):e7573, 2018.
[3] L. Sun *et al*. Effective polyploidy causes phenotypic delay and influences bacterial evolvability. *PLoS Biology*, 16(2):e2004644, 2018.
[4] P. Wang *et al*. Robust growth of escherichia coli. *Current Biology*, 20(12):1099–1103, 2010.
[5] F. Kawano *et al*. A photoactivatable Cre-loxP recombination system for optogenetic genome engineering. *Nature Chemical Biology*, 12(12):1059–1064, 2016.

Single-cell bacterial electrophysiology reveals mechanisms of stress induced damage

Ekaterina Krasnopeeva[1], Chien-Jung Lo[2] and Teuta Pilizota[1]

***Abstract*— Proton motive force (PMF) is at the basis of bacterial energetics. It powers vital cellular processes and defines the physiological state of the cell. Here we use an electric circuit analogy of an *Escherichia coli* cell to mathematically describe the relationship between bacterial PMF, electric properties of the cell membrane and catabolism. We combine the analogy with the use of bacterial flagellar motor as a single-cell "voltmeter" to measure cellular PMF under external stresses. We find that butanol is an ionophore, and functionally characterize light-induced membrane damage. Our approach coalesces non-invasive and fast single-cell voltmeter with a well-defined mathematical framework to enable quantitative bacterial electrophysiology.**

***Index Terms*— bacterial energetics, proton motive force, membrane damage, indole, butanol, photodamage**

I. Motivation

TO stay alive bacteria, like other cells, maintain adequate supplies of free energy, and under various external stresses attempt to stay viable by distributing it to processes essential for coping with the challenge, while simultaneously maintaining core cellular functions. The two main sources of free energy in living cells are adenosine triphosphate (ATP) molecule and proton motive force (PMF). Chemiosmotic theory states that PMF is an electrochemical gradient of protons across the membrane that powers the production of ATP [1]. The ability to measure PMF in bacteria opens a range of currently inaccessible questions that are at the basis of bacterial free energy maintenance, and consequently, survival.

II. Electrical circuit model of the cell

To create a mathematical framework describing the relationship between the PMF and bacterial physiology parameters we represent a cell as an electrical circuit [2]. In this analogy, proton fluxes are currents, oxidative or substrate-level phosphorylation can be considered as an imperfect battery with non-zero internal resistance, and the membrane resistance and capacitance are connected in parallel.

For a cell placed in an external environment whose pH matches cytoplasmic pH, a contribution of the ΔpH to the total PMF is negligible and, thus, PMF becomes equivalent to the membrane voltage, or the drop of potential on the external resistance in the circuit.

Circuit model helps us relate the changes in the physiological state of the cell to the PMF in a variety of stressful environments and uncover the mechanisms through which these stresses induce cellular damage.

III. Single-cell "voltmeter"

We use bacterial flagellar motor (BFM) as an equivalent of a "voltmeter" as the cells are exposed to various external stresses. Flagellar motor speed varies linearly with PMF [3] allowing us to measure relative PMF changes by observing the changes in its rotational frequency.

IV. Conclusion

Using the electric circuit analogy for the membrane fluxes, and BFM as the cell's "voltmeter" we demonstrate the effect of three different stresses on the cell's membrane conductance. We test our approach by testing a known stress, indole treatment, and confirming it is an ionophore [4]. We then show that butanol acts as an ionophore as well increasing membrane conductance linearly with concentration. Finally, we demonstrate that light of the shorter wavelength immediately affects bacterial membrane causing the loss of proton motive force in a power-dependent manner and quantitatively describe the nature of the photodamage. Our approach of combining high-precision PMF measurements and the "electrical circuit interpretation" of the cell serves as a powerful tool needed for quantitative bacterial electrophysiology [5].

This work was funded by Human Frontier Science Program Grant (RGP0041/2015).

[1]Centre for Synthetic and Systems Biology, Institute of Cell Biology, School of Biological Sciences, University of Edinburgh, Alexander Crum Brown Road, EH9 3FF, Edinburgh, UK. Email: e.krasnopeeva@ed.ac.uk

[2]Department of Physics and Graduate Institute of Biophysics, National Central University, Jhongli, Taiwan 32001, ROC. Email: cjlo@phy.ncu.edu.tw

References

[1] P. Mitchell, "Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism," *Nature*, vol. 191, pp. 144–148, 1961.

[2] B. J. Van Rotterdam, W. Crielaard, I. H. Van Stokkum, K. J. Hellingwerf, and H. V. Westerhoff, "Simplicity in complexity: The photosynthetic reaction center performs as a simple 0.2 V battery," *FEBS Letters*, vol. 510, pp. 105–107, 2002.

[3] C. V. Gabel and H. C. Berg, "The speed of the flagellar rotary motor of Escherichia coli varies linearly with protonmotive force." *PNAS*, vol. 100, pp. 8748–8751, 2003.

[4] C. Chimerel, C. M. Field, S. Piñero-Fernandez, U. F. Keyser, and D. K. Summers, "Indole prevents Escherichia coli cell division by modulating membrane potential," *BBA - Biomembranes*, vol. 1818, pp. 1590–1594, 2012.

[5] E. Krasnopeeva, C.-J. Lo, and T. Pilizota, "Single-cell bacterial electrophysiology reveals mechanisms of stress-induced damage," *arXiv:1809.05306*, 2018.

Influence of measurement bias on the interpretation of RNA sequencing results.

Paweł Kuś[1], Roman Jaksik[1], Marek Kimmel[12]

***Abstract*— RNA sequencing is currently one of leading methods used for the analysis of changes in gene expression profile. However, despite its usefulness, numerous reports identified several features related to gene structure which can bias the results, affecting their interpretation. This research investigates some of these factors, as well as normalization methods which might serve to reduce their effect. Using a publicly available RNA sequencing dataset, based on synthetic RNAs added in known concentrations, we identified GC-content as the main source of bias and made an effort to reduce its effect using various normalization tools and highlighting potential improvement strategies.**

***Index Terms*— RNA sequencing, measurement inaccuracies, normalization, GC-content, gene expression profiling.**

I. Introduction

RNA sequencing is, next to oligonucleotide microarrays, the main high throughput method used for the investigation of changes in gene expression profile. Although numerous methods of RNA sequencing have been developed, most of them involve complicated material preparation procedures, including cDNA synthesis and amplification. As many works have reported, gene structure features affect the course of those processes, resulting in biases that can lead to false conclusions in the subsequent analysis, e.g. investigation of regulational dependencies between genes, or identification of molecular markers specific for different types of cells. Most commonly described source of such problems are differences in the nucleotide composition of genes which affect efficiency of PCR amplification [1] and reversed trancription reaction [2], leading to significant underrepresentation of GC-rich genes in the prepared library.

Main goal of the research is to investigate impact of 1) factors that affects results of RNA sequencing experiments and 2) normalization methods that take identified factors into account.

II. Materials and Methods

RNA sequencing data with spike-in material that served as the control group of genes with known expression levels was obtained from NCBI database (ID: GSE49712). Reads were mapped to GRCh38 genome using *STAR* program [3] and low-expressed genes (below 1 TPM) were removed, as indistinguishable from the noise [4]. Gene-level coverage was computed using *coverage* program from the *bedtools* toolset. Impact of chosen factors has been investigated using *biasPlot* method from the *EDAseq* package (gene GC-content and length), *FastQC* tool (GC-content of reads, sequence duplication levels), and *RSeQC* package (RNA degradation level). *EDAseq* methods have been used to reduce GC-content-associated bias. Results were evaluated by identification of differentially expressed genes (DEG), using *edgeR* package, and assessment of AUC values for the ROC curves.

III. Results

Our analysis showed strong association between gene GC-content/length and its coverage. In contrast to length influence, which seemed to have possible biological ground, GC-related effect turned out to depend on the library preparation method, leading to batch effect. The impact of sequence duplication level or RNA-degradation associated effects were negligible for DEG detection. Although the results of EDAseq normalization used to reduce GC-content bias were ambiguous, differences shown in GC-content distribution among reads and in coverage-GC-content relation indicate that there is a need to utilize GC-based correction.

IV. Conclusions

GC-content-associated batch effect was identified as dominant among the analyzed factors, however its correction proved to be non-trivial. For this reason the development of novel normalization strategies, that would address this issue, is needed to further increase the precision of RNA-based gene expression studies.

References

[1] R. Jaksik, M. Iwanaszko, J. Rzeszowska-Wolny, and M. Kimmel, "Microarray experiments and factors which affect their reliability," in *Biology Direct*, 2015.

[2] E. Sendler, G. D. Johnson, and S. A. Krawetz, "Local and global factors affecting rna sequencing analysis," *Analytical Biochemistry*, 2011.

[3] A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T. R. Gingeras, "Star: ultrafast universal rna-seq aligner," *Bioinformatics*, 2013.

[4] S. Harati, J. H. Phan, and M. D. Wang, "Investigation of factors affecting rna-seq gene expression calls," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.

This work was co-financed by Polish National Science Centre grant 2016/23/D/ST7/03665 and by the European Union through the European Social Fund (grant POWR.03.02.00-00-I029). Calculations were carried out using the computer cluster Ziemowit funded by the Silesian BIO-FARMA project No. POIG.02.01.00-00-166/08.

[1]Institute of Automatic Control, Silesian University of Technology, Gliwice, Poland.

[2]Department of Statistics, Rice University, Houston, TX USA.

Contact: pawel.kus@polsl.pl, roman.jaksik@polsl.pl, kimmel@rice.edu

Model for transitions between life and death

Diederik S. Laman Trip[1,2,3] and Hyun Youk[1,2]

***Abstract*— Populations of replicating cells offer the field of biophysics an experimental model system to theoretically study how particle systems transition from out-of-equilibrium (life, replicating) to equilibrium (death, not-replicating). Complemented by experimental observations with yeast we model *in silico* how a population of autonomous, replicating cells goes extinct. This transition from life to death gives rise to physically interesting phenomena, such as fractals and power-law population characteristics. Fractals emerge as statistical self-similarity from the single cell to population scale. A power-law appears as we follow the number of living cells over time, predicting populations that are able to live forever.**

***Index Terms*— phase transition, out-of-equilibrium, fractal, power-law, death process**

MOTIVATION

Physiology of microorganisms can broadly be classified as replicating (life, out-of-equilibrium) or not-replicating (death, equilibrium). Although the goal of every microorganism is replication - one cell dividing into two cells - the ultimate destiny for any cell is death. This transition from a population of replicating cells to an extinct population is poorly understood. Moreover, the conventionally studied existing models such as birth-death or branching processes are in their simplicity unable to capture the complexity observed in biological systems.

MODEL

In this work we study the transition from life to death using a theoretical model complemented with experimental observations. We propose a three-state discrete-time Markov process to gain theoretical understanding of this transition on both a single cell and population level. The states represent living, dormant and dead cells. Our model allows for a heterogenous population, where cell-to-cell variability is modeled by the transition probabilities between the states. This distinguishes our model from conventional analysis by permitting populations in which every cell is unique. Hence, our approach enables us to model complex biological processes while allowing stochastic and deterministic analysis of the system. We complement our theoretical results experimentally with a quantitative study of the extinction of a population of the yeast *S. cerevisiae* in dire conditions, such as nutrient limitation, high temperatures and other life-threatening situations.

RESULTS

Our model matches experimental observations and predicts physically interesting phenomena. First, we explore a branching process interpretation of our model that yields fractal-like growth in conditions where the death rate is close to, but smaller than, the replication rate. Here, the fractals emerge as statistical self-similarity from the single cell to population level. Moreover, we observe these fractals experimentally as highly irregular, fractal-like spatial organisation of yeast grown on solid media. Next, we study the effect of cellar heterogeneity on both the single cell and population level. We explore this feature of the model by assuming that every living cell in the population has an unique probability to die per unit time. Our Markov model predicts that the number of living cells over time decays as a power-law, a phenomenon we observe experimentally as a power-tail. Together, these results suggest that biological systems might not necessarily obey the central limit theorem: statistical averaging, a concept that one is intuitively familiar with, breaks down when assuming that every cell is unique, an assumption that is biologically sensible. Moreover, our observation of a power-law in the most extreme case predicts a cell-population that remains viable forever.

CONCLUSION

We propose a Markov process to model the transition from a population of replicating cells to an extinct population. This theoretical model matches experimental observations. Moreover, both our model and experimental observations suggest that the transition from out-of-equilibrium to equilibrium in a biological system is nontrivial and counterintuitive. Finally, the observation of a power-law in a biological system is conceptually intriguing, as power-laws are typically phenomena akin to phase transitions.

REFERENCES

[1] S. C. R. Clauset A and N. M. E. J, "Power-law distributions in empirical data," *SIAM Review*, vol. 51 (4): 661-703, 2009.

[2] D. Sornette, "Critical phenomena in natural sciences: Chaos, fractals, self-organization and disorder: Concepts and tools," *Springer Series in Synergetics (2nd ed.)*, 2006.

[1]Kavli Institute of Nanoscience, Delft University of Technology, Delft, the Netherlands.

[2]Department of Bionanoscience, Delft University of Technology, Delft 2629HZ, the Netherlands.

[3]Correspondence: d.s.lamantrip@tudelft.nl

Acknowledgments: This work was funded by Netherlands Organisation of Scientific Research (NWO) and the European Research Council (ERC).

Emergence of antibiotic resistance under different environmental dynamics

Ariel M. Langevin[1,2,*], Imane El Meouche[1,2], and Mary J. Dunlop[1,2]

***Short Abstract* — Antibiotic resistance has become a major public health concern as bacteria evolve and adapt to evade drugs, leading to recurring infections and overuse of antibiotics. Previously, we found that bacterial fitness varied depending on the rate of antibiotic introduction. Here, we explore how different introduction rates impact the evolution of bacterial resistance. We study how the antibiotic dynamics impact homogenous and heterogenous bacterial populations experimentally, and simulate different antibiotic treatment dynamics to understand how we can mitigate the emergence of resistance.**

I. Introduction

ALTHOUGH antibiotics have been a major medical feat for the twentieth century, over time the number of recurrent bacterial infections and antibiotic resistant bacteria found in clinical settings has grown at an exponential rate [1,2]. While intravenously administer antibiotics lead to a rapid increase in antibiotics at the site of infection, orally administered antibiotics produce a slower increase in antibiotic concentration [3]. Our previous work highlighted the significance of studying dynamics of antibiotic introduction, since realistic environments rarely see discrete, rapid switches between the presence and absence of antibiotics [4]. Here, we study the impact that antibiotic introduction dynamics have on fitness; we then extend this to explore the impact on the emergence of drug resistance and the fate of the population after antibiotic exposure.

II. Methods & Results

In our work, we study the impact of how the antibiotic is introduced to a bacterial population alters the fitness, survival and, ultimately, the evolution of resistance.

A. Antibiotic Introduction Rates Influence Fitness and Emergence of Resistance for Single Species

Earlier work elucidated that slower introduction rates provided a disproportionate fitness benefit to a co-cultured population of bacteria [4]. The heterogenous co-cultures were composed of cells containing or lacking antibiotic resistance mechanisms, such as AcrAB-TolC efflux pumps. From this, we hypothesize that these differences in fitness could alter the emergence of resistance through permanent genetic changes. Using a modular turbidostat, the eVOLVER [5], we assess how short-term perturbations in antibiotic introduction rates affect bacterial survival, as well as the evolution of resistance.

B. Response of Bacterial Communities to Different Antibiotic Introduction Rates

Secondly, we study how heterogeneous, bacterial communities respond to antibiotic introduction rates, as bacteria are rarely found in homogenous populations in the real world [6]. Preliminary results have demonstrated that the results of a mixed population are not simply the additive results of single species cultures. This indicates that bacterial communities may also play a role in how resistance emerges.

C. Impact of Dynamic Antibiotic Treatment on Emergence of Resistance

Finally, we explore different antibiotic dynamics and how the rate of change on longer term fluctuations impacts the evolution of resistance. We will present results towards experimentally implementing fast and slow antibiotic oscillations to study the emergence of resistance for more dose-like antibiotic concentrations. In addition, using a stochastic, Moran process [7] informed by earlier experimental work, we aim to predict the frequency of different cell fates based upon simulated treatment schedules.

III. Conclusion

Thus far, our results have demonstrated that slower rates of introduction of antibiotics produce higher frequencies of permanent resistance than faster introduction rates. This has implications for improving the efficacy of clinical treatments, while reducing the frequency at which resistance emerges.

References

[1] Zaman S et al. (2017) A Review on Antibiotic Resistance: Alarm Bells are Ringing. *Cureus* 9(6): e1403.

[2] WHO (2016) Fact sheet - Antibiotic resistance. Accessed at https://www.who.int/en/news-room/fact-sheets/detail/antibiotic-resistance on March 6, 2019.

[3] Levison ME, Levison JH. Pharmacokinetics and pharmacodynamics of antibacterial agents. *Infect Dis Clin North Am*. 2009;23(4):791-815.

[4] Langevin AM, Dunlop MJ (2018) Stress introduction rate alters benefit of AcrAB-TolC efflux pumps. *J Bacteriol* 200:e00525-17.

[5] Wong BG, Mancuso CP, et al. (2018) Precise, automated control of conditions for high-throughput growth of yeast and bacteria with eVOLVER. *Nat Biotechnol*, **36**, 614-623.

[6] Wen X, Langevin AM, Dunlop MJ (2018) Antibiotic export by efflux pumps affects growth of neighboring bacteria. *Sci Reports*, **8**, 15120.

[7] Marrec L, Bitbol A-F (2018) Quantifying the impact of a periodic presence of antimicrobial on resistance evolution in a homogeneous microbial population of fixed size. *J Theor Biol*, 457 : 190-198.

Acknowledgements: This work was funded by NSF grant MCB 1347635.

[1]Department of Biomedical Engineering, Boston University, Boston, MA.
[2]Biological Design Center, Boston University, Boston, MA.
[*]E-mail: amlangev@bu.edu

Dynamic fluctuations within an epigenetic landscape underlie gene expression variability

Ryan Lannan[1,2], Alok Maity[1,2], and Roy Wollman[1,2]

Abstract — **Single cell measurements demonstrate that messenger RNA and protein levels can vary wildly from cell to cell. This phenomenon is often attributed to inherent stochasticity in transcription events that can be observed due to the small number of molecules involved. Interestingly, these transcriptional bursts must occur on a fast timescale of few hours. However, the fast time-scale conflicts with reports of slow temporal fluctuations in allele-specific variability. To test the current transcriptional bursting model, we investigated whether expression variability on a much slower timescale exist and whether this variability is allele specific in a model cell line k562 using multi-color fluorescent reporter gene assay. Using a pulse-chase assay we show that cells with high expression levels relax to the population average over a period of multiple days and that this relaxation correlates with local epigenetic marks. Overall our results point to need to account for the dynamic nature of epigenetic regulation of chromatin states when analyzing cell to cell expression variability.**

Keywords **— gene expression noise, fluctuation analysis, chromatin environment, chromatin dynamics**

I. Background

Recent reports demonstrate genetically identical populations of mammalian cells exhibiting significant cell-to-cell variability in the amount of protein a gene produces [1]. This allele-specific variability can have drastic phenotypic consequences, creating differences in cellular response that can alter cell fate [2]. These variations are thought to be the result of the typically small number of molecules involved in gene expression, which creates inherent stochasticity leading to diverse outcomes. This assumption has been built into the two-state model of gene expression, which posits simple, memoryless transitions between off and on states of gene expression, characterized by inactive periods followed by bursts of transcript production [3]. This model is being challenged by work demonstrating slow fluctuations in gene expression identity that exceed a cell generation, implying strong memory [4]. However, leading reports still utilize simple two-state assumptions to model expression in mammalian cells [5].

What is required is a model explaining the breadth of allele-specific variability while providing a mechanism that accounts for the slow fluctuations in expression identity. Our work demonstrates slow temporal fluctuations of gene expression variability, and models the possibility of multiple timescales mediating the decay of expression identity. We illustrate these timescales using fluctuation analysis, including the existence of a cis-mediated, slow fluctuating component, and show that this slow fluctuating component matches dynamic changes in histone modifications.

II. Results

We built a dual reporter system into the K562 cell line utilizing the same promoter and regulatory structure. This design allows for the removal of shared noise with the intent of studying noise intrinsic to one reporter, i.e. cis-effects. We imaged our system for multiple days, revealing a multi-day timescale for which individual cells reached the different states found in a snapshot of a cell population. We developed a multi-timescale model of allele-specific variability, mixing fast, slow and very slow timescales of expression decay.

We separated a system population into three distinct expression states using modified FACS gating to remove the influence of shared noise between our reporters. The populations represent the bottom, middle and top quintile of expression for one of our reporters given conditioning on the second reporter. We then tracked these populations for two weeks, monitoring the relaxation of these distinct states back to the population mean. This work demonstrated a predicted cis-mediated slow fluctuating component. We performed an analogous separation on a single reporter population and performed ChIP-qPCR upon initial separation and after ten days of relaxation. When amplifying from the promoter region of our construct, we found early enrichment of H3K4me3 in the high expression population that dissipates after the relaxation period.

III. Conclusion

This work demonstrates shared behavior between the histone modification H3K4me3 and allele-specific variability within our system. It also raises the possibility of a multi-state model of expression variability in which slow chromatin fluctuations alter transcriptional bursting parameters.

References

[1] Blake WJ, Kaern M, Cantor CR, Collins JJ (2003) "Noise in eukaryotic gene expression," *Nature* 422: 633-637

[2] Paek AL, Liu JC, Loewer A, Forrester WC, Lahav G (2016) "Cell-to-cell variation in p53 dynamics leads to fractional killing," *Cell.* 165 (3): 631-42

[3] Munsky B, Neuert G, van Oudenaarden A, (2012) "Using gene expression noise to understand gene regulation," *Science* 336 (6078): 183-87

[4] Shaffer SM, Raj A, (2018) "Memory sequencing reveals heritable single cell gene expression programs associated with distinct cellular behaviors," *BioRxiv*

[5] Hagemann-Jensen M, Sandberg R. (2019) "Genomic encoding of transcriptional burst kinetics," Nature 565 (7738): 251-54

[1]Department of Chemistry and Biochemistry, UCLA.
[2]Institute for Quantitative and Computational Biology, UCLA.

Constructing Mitochondrial Shape Space

Greyson R. Lewis[1] and Wallace F. Marshall[2]

***Short Abstract* — Mitochondrial networks exhibit complex, dynamic morphologies. Here, I compare experimentally measured morphologies to a variety of computationally simulated models for network construction and evolution, towards understanding the relationship between form and function in organelles.**

***Keywords* — Mitochondria, Bioenergetics, Graph Theory, Topology, Spatial Networks, Temporal Networks, Spatiotemporal Networks.**

I. Extended Abstract

IN many organisms, mitochondria undergo frequent fission and fusion to form dynamic networks [1]. In budding yeast, fission and outer-membrane fusion are regulated by the proteins Dnm1p and Fzo1p, respectively. DNM1 knockout strains, with abnormally highly interconnected networks, exhibit extended lifespan [2]. FZO1 knockout strains, however, exhibit complete network fragmentation and rapidly lose their mitochondrial genomes [3].

Given that mitochondrial fission and fusion levels affect both network morphology and cellular fitness, is there a functional relationship between the two? Here, I aim to characterize mitochondrial network morphology using mathematical topology, computational simulation, and live-cell imaging, toward understanding its relationship with cellular function.

Mathematics has developed a large toolkit for studying networks and graphs. These tools have been applied to many types of well-characterized networks, some built randomly, and some built to be optimal for functions such as maximal interconnectedness [4]. Do mitochondrial networks correspond to well-understood networks from mathematics?

To answer this question, I will use techniques from network theory, graph theory, and applied topology to classify well-understood graph and models, as well as evaluate their similarity to simulated and imaged mitochondrial networks.

References

[1] Nunnari J, et al. (1997) Mitochondrial transmission during mating in Saccharomyces cerevisiae is determined by mitochondrial fusion and fission and the intramitochondrial segregation of mitochondrial DNA. *Mol. Biol. Cell* **8**, 1233-1242.

[2] Scheckhuber CQ, et al. (2007) Reducing mitochondrial fission results in increased life span and fitness of two fungal ageing models *Nat. Cell Biol.* **9** 99-105.

[3] Jones BA, Fangman WL (1992) Mitochondrial DNA maintenance in yeast requires a protein containing a region related to the GTP-binding domain of dynamin. *Genes and Dev.* **6** 380-389.

[4] Holme P, Saramaki J (2012) Temporal Networks. *Phys. Rep.* **519** 97-125.

Acknowledgements: This material is based upon work supported by the National Science Foundation under grant number 1650113, NIH grant 5T32GM008284-27, and the Center for Cellular Construction, an NSF Science and Technology Center.

[1]Department of Biophysics and Biochemistry, University of California San Francisco, San Francisco, CA. E-mail: greyson.lewis@ucsf.edu

[2]Department of Biophysics and Biochemistry, University of California San Francisco, San Francisco, CA. E-mail: wallace.marshall@ucsf.edu

Analytical behavior and comparison between second order genetic regulatory systems

Juan C. Linares[1], and Juan M. Pedraza[2]

The aim of the project was to develop a mathematical method to predict dynamical metabolic variability for second order non-linear genetic regulatory systems, i.e. considering two molecule simple binding reactions. The mathematical procedure is based on asymptotic analysis and gives as a result the behavior of the system as a function of its biological parameters. This allows an analytical comparison between different metabolic pathways. We considered five common genetic regulatory networks and compared their noise behavior, information channel capacities and relaxation times.

A considerable amount of useful chemicals such as biofuels and pharmaceuticals have been successfully produced by merging Synthetic Biology and metabolic engineering with industry needs [1], proving the potential of research in genetic regulatory networks (GRNs). More than two decades have already passed since some of those accomplishments, and we still carry with the limitations that come along with being unaware of the mysteries of the *noise* in GRNs. Two main limitations are the need of trial and error experimentation and the wastage of resources in the process of building viable genetic circuits, attenuated by the usage of computer-aided optimization in most cases [2].

Synthetic GRNs are often designed and built experimentally under controlled laboratory conditions, where environmental fluctuations (and dynamic noise propagation) effects could be easily omitted in optimization processes. The preceding situation leads to the lack of robustness of genetic constructs in large bioreactors outside the laboratory, and consequent unfulfillment of the desired requirements for such circuits.

Adopting the solution of those issues as current challenges in Synthetic Biology has directed the attention of several researchers towards the study of sources and propagation of stochasticity in cellular metabolism. Natural GRNs have provided inspiration in the design of dynamic control strategies that increase bioproduction, but the aimed design principles constitute the need of an integration of the latter with model-based metabolic engineering to develop robust and efficient microbial cell factories [3].

In this context a predicament arises: exactly solvable stochastic descriptions represent a small minority of genetic circuits. Some of the exact solutions reported depict induced transcription followed by translation processes, and steady state solutions of feedback loops in protein production (self-repression and self-induction) [4]. The difficulty of solving a dynamical stochastic system by analytical means comes as non-linearity appears in the rates of the random processes involved. To circumvent this complication several studies based on approximations of the systems have been made: small gaussian noise and slow promoter switching between active and inactive states, to name a few [5]. Those approaches (among others [6]) generate approximate steady state and time varying distributions for GRNs, and hence serve as a basis for understanding non-linear genetic systems.

We present a new way of acquiring an analytical solution of a specific case of non-linear GRNs, where second order addition reactions are considered. In the mathematical process that will be shown, the non-linearity yielded by those reactions is decomposed using asymptotic analysis. The dynamical response of the system to an input is captured as a function of its biological parameters, allowing a functional comparison of the noise behavior, information channel capacity and relaxation times between the architecture of different GRNs.

The suggested comparisons are carried in five common GRNs to exemplify and facilitate the implementation of such requirements. Describing the GRNs taken into account: a system where its output protein is regulated by the input-controlled binary state of another protein (phosphorylation for example); two others in which the input molecule binds directly to a protein, or a RNA strand, interfering in the translation of the output; direct interference in the transcription of the output signal; and a Riboswitch system where the input mediates the binding of two RNA strands, which can promote (or inhibit) the production of an output protein.

REFERENCES

[1] Peralta-Yahya PP, Zhang F, del Cardayre SB, Keasling JD (2012) Microbial engineering for the production of advanced biofuels. *Nature* 488:320–328

[2] Burgard, A.P., Pharkya, P., and Maranas, C. D. (2003) Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.* 84, 647−657.

[3] Liu et al, Dynamic metabolic control: towards precision engineering of metabolism, *J Industrial Microbiology and Biotechnology,* 2018.

[4] Kumar, N. (et al.) Exact distributions for stochastic gene expression models with bursting and feedback. *Phys. Rev. Lett*. 113, 268105 (2014).

[5] Johan Paulsson. Models of stochastic gene expression. *Physics of Life Reviews 2* (2005) 157-175. DOI: 10.1016/j.plrev.2005.03.003.

[6] Schnoerr, D., Sanguinetti, G. & Grima, R. Approximation and inference methods for stochastic biochemical kinetics: a tutorial review. *J. Phys.* A 50, 093001 (2017).

[1]Department of Physics, Los Andes University, Bogotá-Colombia. E-mail: jc.linares147@uniandes.edu.co

[2]Department of physics, Los Andes University. Bogotá-Colombia. E-mail: jmpedraza@uniandes.edu.co

Dynamics of Gene Regulatory Circuits Drive Irreversible State Transitions during Cell Cycle

Ataur Katebi[1], Vivek Kohar[2], and Mingyang Lu[3]

***Short Abstract* — Complex biological processes, such as cell cycle, involve precise cellular state transitions that are controlled by gene regulatory circuits. To evaluate the irreversible state transitions during cell cycle, we applied a systems biology algorithm, named *ra*ndom *ci*rcuit *per*turbation (RACIPE), on yeast cell cycle gene regulatory circuit. We developed statistical analysis algorithms to infer the propensity of state transitions using random models of both stable steady states and limit cycle. We believe the RACIPE-based approach can improve our understanding of the dynamical behavior of gene regulatory circuits.**

***Keywords* — mathematical modeling, gene regulatory circuits, state transitions, cell cycle, and limit cycles.**

I. BACKGROUND

COMPLEX biological processes, such as cell cycle, cell differentiation and circadian rhythm, usually involve precise cellular state transitions that are controlled by gene regulatory circuits. Typical genomic approaches measure data for snapshots of cellular states, but provide little insight into their dynamics. Here, we used cell cycle as a model system to elucidate the cellular state transitions using a systems-biology approach. We applied our recently developed mathematical modeling algorithm, named random circuit perturbation (RACIPE), which models the dynamics of a gene regulatory circuit without precise kinetic parameters [1–3]. RACIPE takes the topology of the gene regulatory circuit as the only input, and generates an ensemble of ODE models with distinct random kinetic parameters. The modeling results from the ensemble of models are then subject to statistical analysis to identify robust feature of the gene regulatory circuit. In this project, we aim to elucidate the irreversible state transitions during cell cycle using RACIPE-generated random models.

II. RESULTS

We applied RACIPE to the core cell cycle gene regulatory circuit of 15 genes and found that the ensemble of RACIPE models consists of both stable steady states and oscillatory states. The steady states cluster into six distinct groups that can be associated with various cell cycle phases. While the time trajectories from oscillatory dynamics travel through the gene expression space of several cell cycle phases. To evaluate the state transitions along a limit cycle trajectory, we trained a neural network model on the RACIPE-generated stable steady states and deployed it to infer the transitions of cell cycle phases along a limit cycle. We found that the limit cycles spanning four or more cell cycle phases are predominantly unidirectional along the cell cycle direction. However, the limit cycles spanning three or fewer phases exhibit random directions. Our results suggest that the gene regulatory circuits not only determine gene expression patterns in different phases but control the directionality of the state transitions within oscillatory dynamics as well.

Next, we explored whether the direction of state transition can be inferred using the stable steady states. We developed a correlation-based method to predict the propensity of transition between two cell cycle phases using the RACIPE-generated simulated gene expression data and the circuit topology. From this approach, we can predict the right direction of state transitions for the cell cycle circuit.

Lastly, we examined possible mechanisms of cell cycle progression by varying the signaling states of the gene circuit. We applied various simulation scenarios to perturb the parameters starting from the models belonging to each of the cell cycle phases and evaluated how those models diffuse to other phases. We found that irreversible transitions can be achieved by either relaxing the parameters toward the global average of all random models (therefore more balanced parameters) or relaxing them towards the parameters of limit cycle models. Our perturbation simulations provide new insight into the possible roles of signaling in irreversible state transitions of cell cycle.

III. CONCLUSION

These results suggest that the ensemble-based modeling approach and the related statistical data analysis framework elucidates the cell cycle dynamics and the cellular state transitions. The approach is applicable to gene regulatory circuits of other biological processes.

REFERENCES

[1] Huang B, Lu M, Jia D, et al. (2017) Interrogating the topological robustness of gene regulatory circuits by randomization, *PLOS Comput. Biol.*, 13(3), pp. e1005456.

[2] Huang B, Jia D, Feng J, et al. (2018) RACIPE: a computational tool for modeling gene regulatory circuits using randomization, *BMC Systems Biology*, 12:74.

[3] Kohar V and Lu M (2018) Role of noise and parameter variation in the dynamics of gene regulatory circuits, npj Systems Biology and Applications, 4: 40.

Acknowledgements: This work was funded by NCI award P30CA034196 and the NIGMS award R35GM128717

The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609

[1]Email: Ataur.Katebi@jax.org, [2]Vivek.Kohar@jax.org, [3]Mingyang.Lu@jax.org

Clocks, Anticipation, and Growth in Bacteria

Michele Monti[1,2], Pieter Rein ten Wolde[2], and David K. Lubensky[3]

***Short Abstract* — Circadian rhythms are widespread across all kingdoms of life, and they are frequently assumed to provide an adaptive benefit by allowing organisms to anticipate diel cycles in their environment. Yet it has proven extremely difficult to determine precisely how such anticipation confers a fitness advantage. Here, we use mathematical modeling to address this question for nitrogen-fixing cyanobacteria. By extending recent work on growth laws in *E. coli*, we show that clock-driven anticipation can allow these organisms to circumvent constraints associated with slow proteome relaxation and thereby to increase their average growth rate.**

***Keywords* — Circadian clock, cyanobacteria, growth rate, fitness, proteome sectors**

I. Background

MANY organisms reorganize their proteome in a circadian fashion in response to the daily nutrient changes in their environment. A striking example is provided by cyanobacteria, which perform photosynthesis during the day to fix carbon. These organisms not only face the challenge of rewiring their proteome every 12 hours, but also the necessity of storing the fixed carbon in the form of glycogen to fuel processes during the night.

In this contribution, we extend the framework developed by Hwa and coworkers for quantifying the relationship between growth and proteome composition to circadian metabolism [1-5]. In this formalism, which has strong experimental support in *E. coli*, the relation between the proteome composition and the growth rate can be quantitatively described by growth laws, which are based on the idea that cells need to balance the supply of amino-acids via catabolic and anabolic reactions with the demand for amino-acids in the synthesis of proteins by ribosomes.

We then apply this framework to investigate the circadian metabolism of the cyanobacterium *Cyanothece*, which fixes not only carbon during the day, but also nitrogen during the night, storing it in the polymer cyanophycin [6]. Our approach differs from previous attempts to quantify the metabolic benefits of circadian clocks [7,8] in that, rather than seeking to reconstruct an entire metabolic network, it takes a coarse-grained, phenomenological approach. This perspective allows us to clearly identify the fundamental physical origin of different observed behaviors in our model.

Cyanobacteria remain the only organisms for which there is direct experimental proof that circadian clocks confer a fitness benefit (albeit in the model strain *S. elongatus*, which does not fix nitrogen) [9,10]. Our calculations lay the groundwork for further work to develop a clear quantitative understanding of the origins of this advantage.

II. Results

Our analysis reveals that the need to store carbon and nitrogen tends to generate an extreme growth strategy, in which the cells predominantly grow during the day, as observed experimentally. This strategy maximizes the growth rate over 24 hours, and can be quantitatively understood by the bacterial growth laws. Our analysis also shows that the slow relaxation of the proteome, arising from the slow growth rate, puts a severe constraint on implementing this optimal strategy. Yet, the capacity to estimate the time of the day, enabled by the circadian clock, makes it possible to anticipate the daily changes in the environment and mount a response ahead of time. This significantly enhances the growth rate by counteracting the detrimental effects of the slow proteome relaxation.

References

[1] Scott, M et al. (2010) Interdependence of cell growth and gene expression: origins and consequences. *Science* **19**, 1099-102.
[2] You, C et al. (2013) Coordination of bacterial proteome with metabolism by cyclic AMP signaling. *Nature* **500**, 301-6.
[3] Huis, S et al. (2015) Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Mol Syst Biol* **11**, 784.
[4] Mori, M et al. (2017) Quantifying the benefit of a proteome reserve in fluctuating environments. *Nat Commun* **8**, 1225.
[5] Erickson, DW et al (2017) A global resource allocation strategy governs growth transition kinetics of Escherichia coli. *Nature* **551**, 119-23.
[6] Reddy, KJ et al. (1993) Unicellular, aerobic nitrogen-fixing cyanobacteria of the genus Cyanothece. *J Bacteriol* **175**, 1284-92.
[7] Hellweger, FL (2010) Resonating circadian clocks enhance fitness in cyanobacteria in silico. *Ecol Modelling* **221**, 1620-29.
[8] Reiners, AM et al. (2017) Cellular trade-offs and optimal resource allocation during cyanobacterial diurnal growth. *PNAS* **114**, E6457-65.
[9] Ouyang, Y et al. (1998) Resonating circadian clocks enhance fitness in cyanobacteria. *PNAS* **95**, 8660-4.
[10] Woelfle MA, et al. (2004) The adaptive value of circadian clocks: an experimental assessment in cyanobacteria. *Curr Biol* **14**, 1481-6.

Acknowledgements: This work was funded in part by NSF grant DMR-1056456 and by the Netherlands Organisation for Scientific Research (NWO).

[1]Centre for Genomic Regulation (CRG), The Barcelona Institute for Science and Technology, Dr. Aiguader 88, 08003 Barcelona, Spain. E-mail: michele.monti@autistici.org
[2]AMOLF, Science Park 104, 1098 XE Amsterdam, Netherlands. E-mail: tenwolde@amolf.nl
[3]Department of Physics, University of Michigan, Ann Arbor, MI 48109-1040, USA. E-mail: dkluben@umich.edu

Single-cell, dynamic interrogation of antibiotic resistance acquisition

Jean-Baptiste Lugagne[1,2], Nathan M. Tague[1,2], and Mary J. Dunlop[1,2]

***Short Abstract* — We are developing an automated experimental platform combining microfluidics, optogenetics and single-cell timelapse microscopy to dissect genetic regulation that facilitates the acquisition of antibiotic resistance in populations of cells tolerant to antibiotics. We propose to drive the AcrAB-TolC multi-drug efflux pump using light signals and monitor the signal propagation to the DNA mismatch repair protein MutS as well as other genes involved in the emergence of antibiotic resistance. Tolerance has been shown to be a stepping stone towards antibiotic resistance, and our goal is to develop a quantitative understanding of the dynamics of this process.**

***Keywords* — Gene Regulatory Networks; Antibiotic Resistance; Antibiotic Tolerance; Real-Time Control; Optogenetics; Microfluidics.**

I. INTRODUCTION

Gene regulatory networks are signal processors, and their response to stimuli can be complex and transient in nature. Monitoring the dynamics of gene expression can in some cases be the only way to detect connections between elements of a genetic system [1]. This has led to the development of novel, automated experimental platforms that can monitor and perturb gene regulatory networks at the single-cell level in real time.

We propose to use this approach to investigate the complex, multi-step process of antibiotics resistance acquisition. Antibiotics tolerance has been shown to be a stepping stone for bacteria to acquire resistance, and recent studies suggest a direct influence of tolerance mechanisms on the emergence of resistant genotypes [2,3]. Our goal is to discover and dissect the dynamic relationships between key genes known to be involved in tolerance and resistant mutation acquisition. We are currently focusing on the interplay between the AcrAB-TolC multi-drug efflux pump, an important element of stress tolerance in *Escherichia coli*, and the MutS mismatch repair protein.

II. EXPERIMENTAL SETUP

We designed and built a so-called "Mother machine" microfluidic device [4], with computer-controlled valves and pumps. This setup, in combination with automated time-lapse microscopy, allows us to track single-cells for over 24 hours, and to dynamically change their chemical environment. In addition to switching between antibiotics or nutrients levels, this setup makes it possible to use time-varying profiles of chemical inducers to perturb, and possibly steer, gene expression dynamically.

In parallel, we are developing optogenetic tools to perturb and drive the expression level of genes of interest within their physiological range. We are working on both optogenetically-driven CRISPRi/a perturbations and chromosomal insertion of the CcaS/R [5] optogenetic system to be able to take over expression of the *acrAB* operon, and plan to extend this procedure to other genes of interest. We are also working on on-line image segmentation and cell tracking, in an effort to close the loop between automated sensing and actuation and to be able to perform real-time control of gene expression and thus be able to set arbitrary gene expression profiles.

III. PERSPECTIVES

By changing both the antibiotic concentrations the cells are subjected to over time and setting AcrAB-TolC levels, we will obtain rich information on the interactions between the efflux pump tolerance mechanism and the MutS-driven mutation strategy of the cells. We propose to then work toward a theoretical model of tolerance to resistance acquisition that describes the mesh of interactions between AcrAB-TolC, MutS and other genes involved in the evolution of antibiotic resistance.

REFERENCES

[1] Lugagne, J.-B., & Dunlop, M. J. (2019). Cell-machine interfaces for characterizing gene regulatory network dynamics. *Curr Opin Syst Biol*.

[2] Levin-Reisman, I., Ronin, I., Gefen, O., Braniss, et al. (2017). Antibiotic tolerance facilitates the evolution of resistance. *Science.*

[3] El Meouche, I., & Dunlop, M. J. (2018). Heterogeneity in efflux pump expression predisposes antibiotic-resistant cells to mutation. *Science.*

[4] Wang, P., Robert, L., Pelletier, J., Dang, W. L., et al. (2010). Robust Growth of Escherichia coli. *Curr Biol*.

[5] Ong, N. T. X., & Tabor, J. J. (2018). A miniaturized E. coli green light sensor with high dynamic range. *ChemBioChem*

Restarting life on demand: Distinguishing dormancy from death by resuming life in yeast

Théo Maire[1,2], Tim Allertz[1,2], Max Betjes[1,2] and Hyun Youk[1,2]

***Abstract*— Dormant cells show no outward signs of life. A natural question is how then one can distinguish them from dead cells. We addressed this question by using glucose to wake-up yeast spores and induce them to enter replicating state (i.e., germinate). We discovered that only a fraction of spores germinate for a given glucose concentration. By examining why some spores do not germinate despite having ample glucose, we discovered that we can quantify both the spore's ability to germinate and how close it is to death by measuring its ability to produce proteins.**

***Index Terms*— dormant cells, Yeast spores, germination, gene regulation, decay to death.**

I. Background

Stopping life indefinitely and then resuming it at the press of a button is an idea that has captivated movie makers, science fiction writers, and the general public. An open question is whether one can indeed completely stop and then resume, after many years, the life of any animals, including humans. Yeast spores are ideal for investigating the halting and resuming of cellular life. Yeast spores do not outwardly appear to be living. They neither move nor divide, exist without any external energy sources, and are believed to maintain faint, if any, intracellular dynamics. While we know how yeasts form spores when they are starved of nutrients and how a newly introduced energy source (glucose) wakes-up the spores and cause them to re-enter replicative life, little is known about the processes that occur, if any, inside the spores during dormancy before glucose is added, how long the dormancy can last (and what determines this timeline), and why some spores cannot wake up (thus considered dead) after a long enough time without any nutrients. Why a dormant spore dies, and how one would distinguish a dormant spore from a dead spore, when both show no outward signs of life, are open questions that we sought to explore.

II. Results

We explored these questions by investigating how glucose, the necessary energy source, germinates yeast spores (i.e., restart cell replication). Using time-lapse microscopy, we first measured how many spores germinated after we added a fixed concentration of glucose. At different glucose concentrations, we found that only some of the spores germinated and that the average time taken to germinate is only dependent on the glucose concentration.

We then investigated what happened to the spores that did not germinate. After no more spores germinated after encountering ample yet less than a saturating concentration of glucose (2%), we added more glucose until the glucose concentration saturated and found that all un-germinated spores germinated. More in depth analyses (including genome-wide RNA-seq) indicated that un-germinated spores showed signs of progression towards germination after encountering the first pulse of glucose. Importantly, we found that the un-germinated spores were primed so that they germinated faster upon encountering more glucose.

We then explored why some spores failed to germinate while others did. By inducing GFP production and giving different glucose concentrations to spores, we found that spores with a lesser ability to produce proteins - indicated by lesser production of GFP - required greater amount of glucose to germinate. By further analyses, we could determine that (1) dormant yeast spores are "programmed" to germinate only above a specific amount of glucose, (2) this amount varies between genetically identical spores and (3) this variability relates to the ability to produce proteins which we could quantify using GFP.

Finally, by identifying dead spores to be ones that even a saturating amount of glucose cannot germinate, we asked how our findings may explain how and why a dormant spore dies. We triggered the spores to die by incubating them in water for an extended period of time (up to months). While the fraction of spores found to be dead increased over time, the remaining dormant spores had their ability to produce proteins reduced. This indicated that dormant spores gradually lose their ability to make proteins until they reach a point of no return, at which point they become unable to germinate (i.e. dead).

III. Conclusions

Our results suggest that being dormant is not a single state but rather a continuous spectrum, with active replication and death at the extreme ends. Interestingly, this spectrum can be quantified by the spore's ability to produce proteins. Our results also reveal a way of quantifying how close a cell is to death without using its chronological or replicative age as a marker.

This work was funded by ERC Starting Grant, NWO Vidi, EMBO Young Investigator Program, and CIFAR Global Scholars Program.

[1]Department of Bionanscience,
[2]Kavli Institute of Nanoscience
Delft University of Technology, Delft 2628, the Netherlands. Email: t.t.h.maire@tudelft.nl

Minimal Boolean model of biological development

Somya Mani[1] and Tsvi Tlusty[2]

***Abstract*— All known multicellular organisms develop from single celled zygotes, be it simple creatures like *Volvox* with two cell-types, or complex creatures like humans with more than 200 cell-types. We seek to understand how biological development generates such diversity. To answer this question, we have developed a minimal Boolean model of development. For any given initial cell, the model outputs the cell-type lineage map of the full multicellular organism. Our results allow us to comment on the forms of cell-type lineage maps we should expect for real multicellular organisms, given our current understanding of the minimal machinery of development.**

***Index Terms*— Development, genome regulation, asymmetric cell-division, cell signaling, cell-type lineage map, graph topology**

I. Introduction

During development, cells of an embryo divide, specialize and gain functionality. This series of divisions and specializations is often represented as a cell-type lineage map; a graph whose nodes represent cell-types, and edges represent mother-daughter relationships between these cell-types. Across all organisms, two processes drive this specialization: asymmetric cell division and cell signaling. Asymmetric cell division is essential for generating cell-type diversity, especially in the initial stages of embryo development. While signals received by cells control gene expression, and thereby control cell-type [1]. But we don't understand if this minimal description of development is sufficient to explain the diversity of extant multicellular life. Also, cell-type lineage maps are widely believed to resemble binary trees; each level of the tree being more specialized than the last. But recent single cell transcriptomics studies have begun to challenge this view [2]. What kind of graphs then represent cell-type lineage maps?

To resolve such questions, we have built a minimal Boolean model of biological development. Although a lot is known about genetic interactions and cell signaling in cases such as the development of wing imaginal discs in *Drosophila* [3], polarity establishment in *C. elegans* [4], etc, we do not include detailed machinery of development. Instead, in order that our results are not constrained to a few model organisms, while we use genome regulation, asymmetric cell division and cell signaling as ingredients of the model, we do not conform to any particular form of these processes. Using our model, we map out the diversity of possible cell-type lineage maps, and link different topologies of lineage maps to qualitative features of asymmetric cell division, cell signaling and genome regulation.

This work was funded by IBS-CSLM.

[1]IBS-CSLM, Ulsan, South Korea. Email: somyamn@gmail.com
[2]IBS-CSLM, Ulsan, South Korea. Email: tsvitlusty@gmail.com

II. Results and conclusions

In our model, genes can be in either of two states: 'on' or 'off'. Therefore, a system with N genes has 2^N cell types. Depending on the underlying gene regulatory dynamics, some of these cell-types are stable, and others lie in their basins. We represent asymmetric cell division and signal transduction using randomly generated adjacency matrices. We find that parameters relating to the sizes of basins of stable cell-types, degree of asymmetry of cell-division, and the number of cell-pairs exchanging signals, are all important in determining lineage map topology. Lineage maps obtained in the model span a range of topologies. We call strongly connected lineage maps (including graphs with single nodes) unicellular [5], and lineage maps that are directed acyclic graphs multicellular. Multicellular graphs include chain-like and tree-like topologies.

In our model, almost all genomes, under some condition of asymmetric division and signaling, give rise to all the above graph topologies (the exceptions are likely due to sampling). More than 50% of the lineage maps found are unicellular. Small chains are the next most widespread topology. This implies that there are many easily accessible ways to make chain-like lineage maps. This is consistent with the fact that the simplest multicellular organisms; *Volvox carterii*(the most recent organism to evolve multicellularity) [6], and some bacteria, such as *Anabaena* [7], have chain-like lineage maps. Surprisingly, tree-like topologies are very rare in our model. This leads us to believe that either trees are not representative of lineage maps, or, there are more essential ingredients to development other than just cell-division and signaling.

References

[1] S. F. Gilbert, "Developmental biology," vol. 10, 2013.
[2] D. E. Wagner, C. Weinreb, Z. M. Collins, J. A. Briggs, S. G. Megason, and A. M. Klein, "Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo," *Science*, vol. 360, no. 6392, pp. 981–987, 2018.
[3] J. V. Beira and R. Paro, "The legacy of drosophila imaginal discs," *Chromosoma*, vol. 125, no. 4, pp. 573–592, 2016.
[4] L. Rose and P. Gönczy, "Polarity establishment, asymmetric division and segregation of fate determinants in early c. elegans embryos," in *WormBook: The Online Review of C. elegans Biology [Internet]*. WormBook, 2005.
[5] P. Greulich, B. D. MacArthur, C. Parigini, J. A. Briggs, and R. J. Sanchez Garcia, "Stability and steady state of complex cooperative systems: a diakoptic approach," *arXiv: 1903.02518*, 2019.
[6] D. L. Kirk, "A twelve-step program for evolving multicellularity and a division of labor," *BioEssays*, vol. 27, no. 3, pp. 299–310, 2005.
[7] J. W. Golden and H.-S. Yoon, "Heterocyst development in anabaena," *Current opinion in microbiology*, vol. 6, no. 6, pp. 557–563, 2003.

Addressable, "Packet-Based" Intercellular Communication through Plasmid Conjugation

John P. Marken[1] and Richard M. Murray[2]

***Short Abstract* — We develop a system for implementing "packet-based" intercellular communication in an engineered bacterial population via conjugation. Our system uses gRNA-based identification markers that allow messages to be addressed to specific strains via Cas9-mediated cleavage of messages sent to the wrong recipient, which we show reduces plasmid transfer by four orders of magnitude. Integrase-mediated editing of the address on the message plasmid allows cells to dynamically update the message's recipients *in vivo*. As a proof-of-concept demonstration of our system, we propose a linear path scheme that would propagate a message sequentially through the strains of a population in a defined order.**

***Keywords* — Plasmid Conjugation, Information Transfer, Intercellular Communication, Engineered Populations**

I. Introduction

As engineered multicellular systems become more complex, they will need to accommodate intercellular communication involving messages with higher information content and dimensionality [1]. A packet-based communication system uses a single architecture to transmit messages regardless of the complexity of the messages' information content. Such a system is therefore well-suited for addressing the need for high-capacity intercellular communication.

Conjugation is a process where genetic elements ranging from small transposons to large plasmids are transferred horizontally between spatially-adjacent cells [2,3]. Conjugation can therefore be used as the basis for a biological "packet-based" communication system, as the mechanism of transmission is unaffected by the length of the message. Additionally, previous work has engineered non-transferrable "helper plasmids" which give a cell the ability to transfer any plasmid that encodes the appropriate origin of transfer sequence, thus distinguishing senders from receivers [4].

Onto this existing framework we develop a message addressability system wherein each strain in a population expresses Cas9 and a unique identifying guide RNA (gRNA). Message plasmids contain an "address" region that consists of binding sites for various gRNAs, so that if the plasmid is sent to a recipient whose gRNA matches a site on the address, the plasmid is cleaved by Cas9 and degraded. The binding sites on the address are flanked by integrase attachment sites so that cells can use integrase-mediated insertion, deletion, or exchange to dynamically update the recipient list of their messages *in vivo*.

II. Summary of results

We demonstrate that Cas9-mediated cleavage of a plasmid leads to its degradation in *E. coli*, and that this process occurs orthogonally between different gRNA-binding site pairs. We then demonstrate that when a receiver strain expressing Cas9 and a gRNA are cocultured with a sender strain with a matching binding site on its message plasmid, the steady-state population fraction consisting of transconjugants is reduced by four orders of magnitude when Cas9 is induced. In contrast, Cas9 induction has no significant effect on the transconjugant density when the binding site does not match the gRNA. This demonstrates that our system can orthogonally block plasmid conjugation to specific receiver strains with low levels of spurious degradation due to leak.

By inducing integrase-mediated cassette exchange and repeating the above experiments, we demonstrate that *in vivo* editing of a plasmid's address region successfully alters its degradation profile, i.e. a binding site that was formerly mismatched becomes matched and vice versa, with the expected consequences for blocking plasmid transfer.

In addition to these results we propose the design of a linear path scheme that uses our addressability system to ensure the sequential propagation of a message plasmid through the strains in a population, generalizable to any number of strains.

III. Conclusion

We anticipate that the development of new intercellular communication frameworks will be of great importance in facilitating the engineering of synthetic populations of increasing complexity. Our work provides a system with specific advantages in message addressability and high-dimensional information transmission.

References

[1] Sexton JT and Tabor JJ (2019). Multiplexing cell-cell communication, doi: https://doi.org/10.1101/584664

[2] Firth NE, Ippen-Ihler KA, Skurray RA. Structure and function of the F factor and mechanism of conjugation. Escherichia coli and Salmonella: cellular and molecular biology, 2nd ed. ASM Press, Washington, DC. 1996:2377-401.

[3] Lorenzo-Díaz F, et al. (2017) Crosstalk between vertical and horizontal gene transfer: plasmid replication control by a conjugative relaxase. *Nucleic Acids Research* **45**, 7774-85.

[4] Dimitriu, T, et al. (2014) "Genetic information transfer promotes cooperation in bacteria." *PNAS* **111**, 11103-11108.

[1]Department of Bioengineering, California Institute of Technology. 1200 E California Blvd, Pasadena, CA. E-mail: jmarken@caltech.edu

[2]Departments of Bioengineering and Control and Dynamical Systems, California Institute of Technology. 1200 E California Blvd, Pasadena, CA. E-mail: murray@cds.caltech.edu

This research is supported by the Institute for Collaborative Biotechnologies through cooperative agreement W911NF-19-2-0026 and grant W911NF-09-0001 from the U.S. Army Research Office. The content of the information on this page does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

Physical Stretch Activates Mechanically-Gated Calcium Channels for Nitric Oxide (NO) Formation in the Ectoderm of Chicken Embryos

Adrian Martin,B.S., and Wilfred Denetclaw, Ph.D.

Department of Biology, San Francisco State University, San Francisco, CA 94132

Embryology is the study of fundamental mechanisms that address the origins of tissues and organs and their interactions following fertilization and early development. However, the findings of these mechanistic changes that control gene expression and tissue differentiation are relatively recent. For example, calcium-dependent activation of nitric oxide synthase (NOS) to produce nitric oxide (NO) was recently identified to signal in myogenesis, but how this rise of NO occurs is poorly understood. Therefore, we hypothesize the ectoderm layer under stretch-forces activate mechanically-gated calcium channels to raise calcium (Ca2+) and signal to NOS for NO. To investigate, controlling tension stretch activity was done using HH10 embryos that were egg extracted and pinned to a silicon dish on one side and pulled on the other by a custom made “rake” device. Embryos were imaged using live spinning disc confocal microscopy labeled for Ca2+ using calbryte520 and NO using DAF-2. Stretch-forces induced in the embryo cranial midline of the ectoderm show an immediate two-fold increase of Ca2+ within 3 seconds that lasts for 1 minute before returning to normal levels. However, this strategy, while invoking a Ca2+ rise, did not increase NO. To show the functionality DAF-2 probe and the overall health of the embryo, 1mM ATP added to the ectoderm layer resulted in a substantial increase of NO 30 minutes after treatment. Despite these results, stretch-forces produced a slight rise in Ca2+ and a sharper increase in NO through isolated sheets of ectoderm cultured on a stretchable PDMS membrane which demonstrate the ability of the ectoderm under stretch to generate NO. The embryo stretch may also result in NO elevation after additional modifications of embryo loading with the DAF-2DA probe. In conclusion, these results suggest that the ectoderm possesses the ability to raise NO levels through mechanically-gated calcium channels that signals in myogenesis. NSF STC CCC: 1548297

Multi-Timescale Dynamics of the Cell Cycle-Stress Response Interface

GW McElfresh[1] and Christian Ray[2]

Signaling pathways enable living cells to compute responses to stimuli from the extracellular environment. Growing cells invest energy to grow and divide, thereby diluting previous computations. Because the remnants of previous responses are diluted, an intergenerational memory is imparted: a daughter cell is predisposed to respond in a qualitatively similar manner to its mother cell. Here we analyze the effect of cellular memory and cell cycle on a broad class of bacterial information transfer systems, two-component system modules. Our analysis places non-genetic intergenerational information transfer in a computational context and identifies surprising patterns of cell cycle-stress response interactions.

***Keywords* — two component system, PhoBR, cellular signaling, cell cycle, bacteria, complexity-aware modeling**

I. PURPOSE

CELLULAR signaling pathways allow for cells to respond to environmental stimuli and understanding signaling is critical to understanding cellular physiology. Bacterial response networks that control global gene expression shifts require remarkably low signal fidelity [1,2]. Research seeking to quantify stochasticity in signaling networks typically considers timescales on the order of molecular fluctuations [1,3,4 and others].

We propose that to understand bacterial response networks, multi-order timescales need to be considered. Bacterial growth and gene network activation happen on similar timescales and cannot be readily isolated from each other. However, the half-life of a response protein is around 70% of a cell's lifespan, introducing a generational timescale that becomes relevant in the case of periodic stress signals [5].

We developed a complexity-aware cellular simulation based on PhoBR with a growth model parameterized extensively on experiments [5,6]. The signaling network and downstream regulon are used to calculate an emergent growth rate-based resource partitioning in the cell, allowing for long simulation times to explore the varying timescales of cellular response.

II. RESULTS

We studied responses in the PhoBR two component system to observe the key mechanisms driving the growth in a bacterial stress response.

A. Cellular Memory

Residual stress response proteins increase cellular fitness in the presence of repeated stress response [7]. We quantified the molecular trajectories of the stress response network and found that there are three significant timescales relevant to cellular memory. The sensing protein loses activation within a minute, the regulatory protein maintains activation for several generations, and the downstream stress response proteins are diluted over many generations [5]. Memory is conferred unequally to daughter cells, creating bet-hedging.

B. High-dimensional profiling of signal/cell cycle dynamic interactions

Stochasticity in signaling networks creates heterogenous cell growth trajectories, but a fine-grained picture of the major contributors to growth rate are unclear. We applied dimensional reduction and correlation methods to categorize the leading variables. The downstream regulon is usually the strongest contributor to cellular state due to energetic investment involved in expressing proteins. Other molecular species become transiently correlated with cell cycle effects, suggesting a rich, nonlinear dependence between fast fluctuations and the cell state.

III. CONCLUSION

The physiological context of the cell cycle is an intrinsic driver of signaling that itself is affected by signaling. Studies of signal propagation in living matter must consider this context to gain a complete picture of cellular information processing.

REFERENCES

[1] Cheong *et al*. 2011. "Information transduction capacity of noisy biochemical signaling networks." Science 334 (6054): 354–358.

[2] Suderman, *et al.* 2017. Proceedings of the National Academy of Sciences, 114 (22) 5755-5760; DOI:10.1073/pnas.1615660114

[3] Barato, *et al*. 2014. "Efficiency of cellular information processing." *New Journal of Physics* 16:103024.

[4] Govern, *et al*. 2014. "Optimal resource allocation in cellular sensing systems." *Proceedings of the National Academy of Sciences* 111 (49): 17486–17491.

[5] McElfresh GW & Ray JCJ. 2018. *Intergenerational Cellular Signal Transfer and Erasure.* In: The interplay of thermodynamics and computation in natural and artificial systems. Santa Fe Institute Press.

[6] Marzan, Lolo Wal, & Kazuyuki Shimizu. 2011. Metabolic regulation of *Escherichia coli* and its *phoB* and *phoR* genes knockout mutants under phosphate and nitrogen limitations as well as at acidic condition. *Microbial Cell Factories* 10 (1): 39.

[7] Lambert, Guillaume, & Edo Kussell. 2014. "Memory and fitness optimization of bacteria under fluctuating environments." *PLOS Genetics* 10 (9): e1004556.

[1]Center for Computational Biology, University of Kansas, Lawrence. E-mail: gwmcelfresh@ku.edu

[2]Center for Computational Biology & Department of Molecular Biosciences, University of Kansas, Lawrence. E-mail: jjray@ku.edu

Acknowledgements: This work was funded by NIH grant P20GM103638

Revealing bacterial chromosome organization from Hi-C data using a maximum entropy approach

Joris Messelink[1], Jacqueline Janssen[1,2], and Chase Broedersz[1]

***Short Abstract* — Elucidating the three-dimensional spatial organization of the bacterial chromosome is essential to understand how core genomic processes are spatially regulated inside of the cell. Recent Hi-C chromosome conformation capture experiments provide contact frequency maps of the chromosome, reflecting its highly organized structure. We develop a maximum entropy approach to extract the three-dimensional structure of the bacterial chromosome from such data. Using this approach, we obtain a coarse-grained model for the full distribution of chromosome configurations. We validate the predictive power of our model by experiments on the localization of chromosomal loci in the cell. Our approach opens up a new way of interpreting Hi-C data.**

***Keywords* — Chromosome organization, Hi-C Chromosome Conformation Capture, Maximum entropy modelling**

I. Chromosome organization and Hi-C data

THE bacterial chromosome outsizes the cell by three orders of magnitude, and must thus be highly compacted to fit inside. Importantly, the chromosome needs to not only be condensed, but also organized to facilitate biological functions such as transcription, replication, and segregation.

The high degree of spatial organization of the chromosome has been demonstrated by microscopy experiments [4] and Hi-C chromosome interaction detection techniques [2]. Two striking features from this data include the presence of Chromosomal Interaction Domains (CID's) and a juxtaposed position of the two chromosomal arms. However, extracting the three-dimensional chromosome conformations from this data poses a major challenge.

II. Inferring a maximum entropy model

Our aim is to infer the full distribution of chromosome conformations from Hi-C data on bacteria such as *C. crescentus* and *B. subtilis*, without making any assumptions about the underlying dynamics. To this end, we introduce a new inference method based on a *maximum entropy approach* [3]: we construct a theoretical polymer model that reproduces a set of experimental Hi-C observations, but otherwise imposes as few assumptions as possible.

A. Solving the maximum entropy model

The maximum entropy model can be formally mapped to an equilibrium model of a polymer with interaction energies between pairs of monomers. This enables us to use a range of techniques from statistical physics to efficiently and effectively infer model parameters from experimental data. We demonstrate that the algorithm is able to accurately solve for the interaction parameters for a series of test cases.

B. Independent experimental validation

To validate the model we infer from Hi-C data on *C. crescentus*, we show that the predicted cellular localization of various genes are in accord with experiments [4]. We confirm that this predicted cellular localizations are due to the specific Hi-C interactions inferred from experiment: without using the Hi-C map as input, this spatial organization disappears.

C. Revealing novel features of chromosome organization

Our maximum entropy model reveals novel features of chromosome organization, including local density maps, a characterization of the typical configurations of the chromosome, as well as the average distances between pairs of loci on the chromosome.

III. Conclusion

We develop a maximum entropy model that yields the distribution of chromosome conformations from Hi-C data. Our results are consistent with independent experimental tests, and reveal novel organizational features. Our approach thus opens up a new way of analyzing Hi-C data.

References

[1] Messelink, J., Janssen, J. Broedersz, C. (2019) Maximum entropy model of the bacterial chromosome (In preparation)

[2] Le TB, Imakaev MV, Mirny LA, Laub MT (2013) High-resolution mapping of the spatial organization of a bacterial chromosome. Science *342* (6159), 731-734.

[3] Bialek, W. (2012). *Biophysics: searching for principles*. Princeton University Press.

[4] Viollier, Patrick H., et al. "Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication." *Proceedings of the National Academy of Sciences* 101.25 (2004): 9257-9262.

Acknowledgements: This work was funded by the Graduate School for Quantitative Biosciences Munich (QBM)

[1]Arnold-Sommerfeld-Center for Theoretical Physics and Center for NanoScience,Ludwig-Maximilians-Universität München, D-80333 München, Germany.

[2] Max Planck Institute for the Physics of Complex Systems, D-01187 Dresden, Germany

Decoupling Priming and Desensitization in Response to IFN-alpha Pretreatment

Anusorn Mudla

Abstract

Interferon-alpha (IFN-α) is a major cytokine produced in response to viral infection and clinically important in anti-viral and anti-cancer therapy. Although several key components of the interferon pathway have been characterized, their dynamics in response to repetitive stimulation remain elusive. Here, we systematically studied how IFN-α pretreatment can lead to two contradictory effects: priming and desensitization. We used CRISPR/Cas9 to fluorescently label signal transducer and activator of transcription 1 (*STAT1*) and its downstream gene, interferon regulatory factor 9 (*IRF9*), and monitor their dynamics. We used a microfluidic device to precisely control the durations of the IFN-a pretreatment, break-time and the second stimulation. Single cell quantification from time-lapse microscopy revealed that 2- and 10-hour pretreatment can lead to priming effect while 24-hour pretreatment lead to desensitization both in STAT1 nuclear translocation and the rate of IRF9 induction. To further investigate the mechanism, we knock-downed ubiquitin specific protease 18 (*USP18*), a known negative regulator of IFN-α signaling, using shRNA and found that STAT1 nuclear translocation was restored and the rate of IRF9 induction was significantly higher. Intriguingly, we observed heterogeneity in the desensitization among 24-hour-pretreated cells of which expressed low level of USP18. We developed a mathematical model to describe and predict the effect of pulsatile IFN-α stimulation. As expected, a pulsatile treatment of IFN-α led to higher IRF9 induction compared to a sustained treatment. Our results demonstrate that priming and desensitization of IFN-α is duration dependent and controlled by USP18 as a delayed negative feeback. This discovery provides insight information to improve pharmacokinetic of IFN-α delivery for more effective viral-infected disease and cancer therapy.

Resource reallocation in engineered *Escherichia coli* strains with reduced genomes

Ernesto S. Nakayasu[1,2], Adam Chazin-Gray[1], Deanna L. Auberry[1], Nathalie Munoz Munoz[1], Jeremy D. Zucker[1], Neeraj Kumar[1], Carrie Nicora[1], Hugh D. Mitchell[1], Young-Mo Kim[1], William C. Nelson[1], Robert Egbert[1,3]

***Short Abstract* — Genome remodeling can optimize bioproduction by eliminating non-essential cellular processes. However, a major knowledge gap in optimizing cellular capacity for heterologous expression is how cells reallocate resources when engineered. We investigate this process using a library of sequential deletion strains of *E. coli* W3110 in combination with genetic engineering and multi-omics characterization. We found that amino acid degradation and polyamine synthesis pathways are upregulated as genes associated with large proteome fractions are deleted. These resources can be reclaimed under heterologous expression of a violacein biosynthesis pathway. These findings suggest new approaches to boost productivity, increase evolutionary robustness, and provide biocontainment strategies for engineered organisms.**

***Keywords* — Cellular Capacity Optimization, Synthetic Biology, Genome Engineering, Cellular Resource Allocation, Multi-omics measurements.**

I. Introduction

The bioproduction capacity of organisms is limited by metabolic conversion rates and the allocation of core cellular resources. It has been proposed that deletion of non-essential genes can liberate resources of the cells to increase production of biocommodities. However, a major hurdle for engineering organisms is the lack of knowledge on how cells reallocate resources when metabolic pathways are rewired.

Here we performed proteomics and metabolomics measurements on a collection of *E. coli* sequential deletion variants (KHK library) [1, 2] to elucidate how resources are reallocated when large genomic regions are deleted. Reclamation of these resources was also tested by introducing a genetic burden through overexpression of the *vioABE* protodeoxyviolacein (PDV) reporter system [3].

II. Key Results

A. Genomic deletions increase production of protodeoxyviolacein pigment by the vioABE system. Some sequential deletions of large genome regions lead to increases in pigment production, but others reduce pigment levels and impact growth rate.

B. Resource allocation. Proteomic analysis revealed that the deletion of genes encoding high-abundance proteins induces the degradation of amino acids, which can be reclaimed by inducing cellular burden via overexpressing the *vioABE* system (**Fig. 1A**). Pathway integration with metabolomics data identified glutamine as the central metabolite in amino acid degradation, leading to the production of succinate and polyamines (putrescine and spermidine) (**Fig. 1B**).

C. Production of protodeoxyviolacein is enhanced by glutamine supplementation. Because glutamine was observed to be the central amino acid degraded by the reallocation of cellular resources we hypothesized that supplementing cell cultures with glutamine would enhance PDV production. This hypothesis was confirmed using the *vioABE* reporter pathway (**Fig. 1C**).

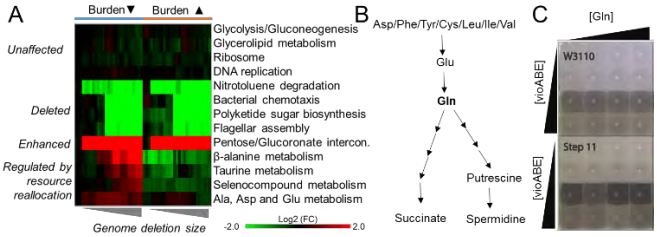


Fig. 1. Resource allocation in the *E. coli* KHK library. (**A**) Pathways regulated in strains with large genomic deletions of non-essential genes. (**B**) Amino acid degradation pathway induced by genome deletions. (**C**) Protodeoxyviolacein production in cultures supplemented with glutamine (Gln).

III. Conclusion

The deletion of highly expressed non-essential genes decreases amino acid demands and cells reallocate these resources to other metabolic processes. These findings open new perspectives for cellular capacity optimization to enhance the productivity of engineered microbes.

IV. References

[1] Mizoguchi, H., Sawano, Y., Kato, J., and Mori, H. (2008) Superpositioning of deletions promotes growth of *Escherichia coli* with a reduced genome, *DNA Res 15*, 277-284.

[2] Hirokawa, Y., Kawano, H., Tanaka-Masuda, K., Nakamura, N., et al. (2013) Genetic manipulations restored the growth fitness of reduced-genome *Escherichia coli*, *J Biosci Bioeng 116*, 52-58.

[3] Balibar, C. J., and Walsh, C. T. (2006) In vitro biosynthesis of violacein from L-tryptophan by the enzymes VioA-E from *Chromobacterium violaceum*, *Biochemistry 45*, 15444-15457.

Acknowledgements: This work was supported by PNNL Laboratory Directed Research & Development funds. Samples were analyzed in the Environmental Molecular Sciences Laboratory, a DOE BER national scientific user facility on the PNNL campus. PNNL is a multi-program national laboratory operated by Battelle for the DOE under contract DE-AC05-76RL01830.

[1]Biological Sciences Division, Pacific Northwest National Laboratory
[2]E-mail: ernesto.nakayasu@pnnl.gov
[2]E-mail: robert.egbert@pnnl.gov

Measurement of noise on added size for *E coli* adder and sizer-like division strategies suggests a multi-step control.

Cesar Augusto Nieto-Acuña[1], Juan Carlos Arias-Castro[1,2],Carlos Arturo Sanchez-Isaza[1,2], Cesar Augusto Vargas-Garcia[3] and Juan Manuel Pedraza[1]

***Abstract*— Recent experimental advances have suggested the Adder mechanism for *Escherichia coli* division control. This means that bacteria grow, on average, a fixed size before division. Here we use new experiments to verify this mechanism with glucose as carbon source and explore deviations of the adder strategy, specifically, the division control of *E. coli* growing with glycerol as carbon source. In this medium, the division strategy is sizer-like, which means that the added size decreases with the size at birth. We propose a mechanistic model to explain our measurements and found that our model can explain not only the slope in the graph of added size vs size at birth but the noise in this relationship. We propose further experiments in order to distinguish between other possible explanations.**

I. Purpose

Bacterial homeostasis, this is, how these cells synchronize their growth with their division-time to control their size in a steady growth, is an important problem in Systems biology. Recently, experimental techniques have been improved enough to allow high throughput dynamics in rod-shape bacteria like *Escherichia coli*[1].

Most of these experiments have suggested the "adder mechanism" of division control, consisting on incorporating a fixed amount of cell wall material before dividing[2], [3]. Sizer-like mechanisms are common deviation of the adder paradigm[4]. These mechanisms occur when the added size just simply decreases with the size at birth. They became important mainly after their observation over yeast[5], slow-growing cells[6] and bacteria growing sub-optimal growth media[7].

We performed single-cell tracking experiments for *E. coli* using a *mother machine* micro-fluid[1]. The main goal of these experiments was to track multiple cell cycles to study the noise in added size during these cycles in two different media: typical M9[8] (which is adder) and M9 with Glycerol as carbon source instead of Glucose (which is sizer-like)[1].

This work was funded by COLCIENCIAS convocatoria doctorados nacionales 647.

[1]Department of Physics, Universidad de los Andes, Bogotá, Colombia. Email: ca.nieto13@uniandes.edu.co

[2]Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA.

[3] Faculty of Mathematics and Engineering, Fundacion universitaria Konrad Lorenz, Bogotá, Colombia.

Currently, there is no clear mechanism behind the division control of these microorganisms but recent experiments[9] and theoretical approaches [10] have revealed some insights suggesting an stochastic process consisting on dividing after reach some threshold by multiple steps. However, these ideas seem to work well in the adder mechanism and is not clear how a sizer-like mechanism could appear.

Here, we propose a mechanistic mechanism which can predict the sizer-like behavior and can fit the noise on added size that we measured in the *mother machine* experiments. This mechanism is based on a splitting rate function which is proportional to a power of the cell size. We derive some analytical formulas and discuss some limitations and possibilities that this framework can give to understand better this phenomenon.

References

[1] S. Taheri-Araghi, S. Bradde, J. T. Sauls, N. S. Hill, P. A. Levin, J. Paulsson, M. Vergassola, and S. Jun, "Cell-size control and homeostasis in bacteria," *Current Biology*, vol. 25, no. 3, pp. 385–391, 2015.

[2] M. Campos, I. V. Surovtsev, S. Kato, A. Paintdakhi, B. Beltran, S. E. Ebmeier, and C. Jacobs-Wagner, "A constant size extension drives bacterial cell size homeostasis," *Cell*, vol. 159, no. 6, pp. 1433–1446, 2014.

[3] S. Jun, F. Si, R. Pugatch, and M. Scott, "Fundamental principles in bacterial physiology—history, recent progress, and the future with focus on cell size control: a review," *Reports on Progress in Physics*, vol. 81, no. 5, p. 056601, 2018.

[4] C. A. Vargas-Garcia, M. Soltani, and A. Singh, "Conditions for cell size homeostasis: A stochastic hybrid system approach," *IEEE Life Sciences Letters*, vol. 2, no. 4, pp. 47–50, 2016.

[5] J.-B. Nobs and S. J. Maerkl, "Long-term single cell analysis of s. pombe on a microfluidic microchemostat array," *PloS one*, vol. 9, no. 4, p. e93466, 2014.

[6] M. Wallden, D. Fange, E. G. Lundius, Ö. Baltekin, and J. Elf, "The synchronization of replication and division cycles in individual e. coli cells," *Cell*, vol. 166, no. 3, pp. 729–739, 2016.

[7] M. Priestman, P. Thomas, B. D. Robertson, and V. Shahrezaei, "Mycobacteria modify their cell size control under sub-optimal carbon sources," *Frontiers in cell and developmental biology*, vol. 5, p. 64, 2017.

[8] C. Harwood and S. Cutting, "Chemically defined growth media and supplements," 1990.

[9] F. Si, G. Le Treut, J. T. Sauls, S. Vadia, P. A. Levin, and S. Jun, "Mechanistic origin of cell-size control and homeostasis in bacteria," *bioRxiv*, p. 478818, 2018.

[10] K. R. Ghusinga, C. A. Vargas-Garcia, and A. Singh, "A mechanistic stochastic framework for regulating bacterial cell division," *Scientific reports*, vol. 6, p. 30229, 2016.

Generating Cell Fate Patterns via Mechanical Stress in Human Stem Cells

Hayden S. Nunley[1], David K. Lubensky[1], Xufeng Xue[2], Jianping Fu[2]

***Short Abstract* — Our current knowledge of fate patterning in the early embryo focuses on diffusible chemical signals, or morphogens. The role of mechanical signals in fate patterning in two-dimensional embryonic tissues remains poorly understood. Recent experiments in stem cell colonies on micropatterned substrates demonstrated that neuroectoderm differentiation, an early patterning event, occurs in the absence of exogenous morphogen gradients [1]. We propose a continuum model for this process in which cell fate is biased by mechanical stress. This model predicts a non-monotonic dependence of the width of the outer fate domain on substrate stiffness. Preliminary experimental results are consistent with model predictions.**

***Keywords* — Neuroectoderm, Neural Plate Border (NPB), Micropatterned Substrate, Micropost Array, Domain Wall**

I. Background

THE proper development of an embryo depends on fate specification events in which a field of initially equivalent cells differentiates in a spatially controlled manner [2]. A key example is the process of neural induction in which a strip of cells in the outer epithelial layer of the vertebrate embryo differentiates into neural cells, forming the neural plate (NP), flanked by the cells of the neural plate border (NPB) [1, 3-4]. Classic studies of this system confirmed the role of diffusible chemical signals from neighboring tissues in patterning [3-4].

Recent experiments have probed this process outside of the embryo [1]. In these experiments, human stem cells bind to a small circular region, coated with protein, on a thin PDMS substrate. The supplied chemical medium that induces neural differentiation is uniform. The colonies differentiate into NP cells surrounded by a ring of NPB cells. After ruling out many endogenous chemical gradients, the authors conclude that the *in vitro* patterning occurs via mechanics [1].

We develop a model of the two-dimensional cell layer in which cell fate determines contractility, and stress biases fate. This model reproduces a key observation, that the concentric width of the NPB domain is approximately independent of colony size. The model predicts that the NPB domain width should depend non-monotonically on substrate stiffness.

II. Results

A. Traction forces reveal fate-contractility correlation

We measured traction forces between the cell layer and the substrate via displacements of microposts. We fit post displacements to a model of a contractile continuum medium [5-6]. These fits reveal a ring of contractile cells, which we identify as the NPB domain. The fits constrain mechanical properties of the layer, and thus our model parameters.

B. Linear coupling of stress to bistable fate reproduces independence of NPB domain size on colony diameter

We model the fate as a bistable variable, w, stable at $w \approx 0$ and $w \approx 1$. We assume that contractility is proportional to w such that $w \approx 1$ is the contractile (NPB) fate. Coupling in-plane pressure to fate, we reproduce the approximate independence of NPB domain width on colony size [1].

C. Prediction of non-monotonic dependence of outer domain width on substrate rigidity

The NPB domain width depends on the ratio of stiffnesses of the cell layer and the substrate. Our model predicts that on glass, a very stiff substrate, the NPB domain only contains the outer-most cells, now confirmed experimentally. The domain width increases as substrate stiffness decreases. For substrates sufficiently softer than the control ($0.5\ kPa/\mu m$), the model predicts that the sign of the effect (i.e., dependence of NPB domain width on stiffness) reverses because the outer cells cannot generate enough stress to bias cells to the NPB fate.

III. Conclusion

The role of mechanics in fate patterning remains poorly understood. This quantitative model of mechanics-guided differentiation generalizes the concept of morphogens from chemical to mechanical, and makes novel predictions about the role of substrate rigidity in patterning of quasi-2D tissues.

References

[1] Xue X, et al. (2018) Mechanics-guided embryonic patterning of neuroectoderm tissue from human pluripotent stem cells. *Nature Materials* **17**, 633-641.

[2] Chan CJ, Heisenberg C-P, Hiiragi, T (2017) Coordination of morphogenesis and cell-fate specification in development. *Curr Biol* **27** R1024-R1035.

[3] Martyn I, et al. (2018) Self-organization of a human organizer by combined Wnt and Nodal signalling. *Nature Materials* **17**, 633-641.

[4] Spemann H, Mangold H (1924) Induction of embryonic primordia by implantation of organizers from a different species. *Roux's Arch Entw Mech* **100**, 599-638. Reprinted and Transl. *Int J Dev Biol* **45**, 13-38.

[5] Edwards C, Schwarz U (2011) Force Localization in Contracting Cell Layers. *Phys Rev Lett* **107**, 128101.

[6] Banerjee S, Marchetti MC (2011) Substrate rigidity deforms and polarizes active gels. *Europhysics Lett* **96**, 28003.

Acknowledgements: The theoretical material is based up work supported by NSF grant DGE-1256260.

[1]Departments of Physics and Biophysics, University of Michigan, Ann Arbor, MI, USA. E-mail: nunley@umich.edu, dkluben@umich.edu

[2]Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI, USA. E-mail: xufeng@umich.edu, jpfu@umich.edu

Cytoplasmic density dynamics in fission yeast

Pascal D Odermatt[1,2], Gabriella Estevam[1], Kerwyn Casey Huang[2] and Fred Chang[1]

***Short Abstract* — We have established an imaging-based approach enabling spatial and temporal quantification of cytoplasmic density in the rod shaped fission yeast to investigate mechanisms coordinating cell growth and biomass accumulation. Our results show cytoplasmic density decreasing during the phase of polar cell growth, and increasing at the end of the cell cycle as growth stops upon mitotic entry, suggesting a continuous biomass accumulation independent from volume expansion. Further, we use mutants and drug treatment to alter cell cycle progression, cell growth patterns and biosynthetic processes to investigate their effect on the regulation of cytoplasmic density.**

***Keywords* — Cytoplasmic density, fission yeast, cell growth**

I. Background

The density of the cytoplasm may be a critical but underappreciated factor for many cellular reactions within the living cell. Altered cytoplasmic density has been linked to chromosomal missegregation[1], senescence[2] or cell mechanical properties[3], however physical principles regulating cytoplasmic density, specifically potential links between cell growth and biomass accumulation, are not fully understood.

Fission yeast, *S. pombe,* is a powerful eukaryotic model for studying the cell cycle and growth due to its simple geometry and growth pattern. Various mutants with perturbed cell cycle progression or altered growth patterns are available, making it a favorable organism to study underlying physical principles linking growth and cytoplasmic density.

II. Results

We have developed the use quantitative phase imaging based on brightfield image stacks[4] as a tool to estimate the cytoplasmic density of individual yeast cells. We estimate the cytoplasmic density of interphase *S. pombe* cells to be approximately 200-250 mg/ml. Fission yeast exhibit tip growth during interphase and cease growth during mitosis and cytokinesis. Measurements of the density dynamics show that cytoplasmic density varies in the cell cycle; it falls slightly in G2 phase and gradually increases during mitosis and cytokinesis reaching a maximum just before division.

To further assess the influence of specific cell cycle stages and growth patterns on the dynamics of the cytoplasmic density we use various mutants and drug treatments. Cells remaining in interphase for an extended period of time exhibiting fast growth are elongated, while cells blocked during cytokinesis stop growing. These mutants allow us to investigate the effects of changes in the growth rate on accumulation of biomass at specific cell cycle stages.

Similarly, chemical inhibition of the secretory pathway causes cells to stop growing while biomass accumulation continues, resulting in increased cytoplasmic density. In contrary, cytoplasmic density does not increase when growth is altered due to inhibition of ribosomal protein synthesis pathways.

Perturbation of cytoplasmic density can additionally have consequences for single cell in regards to cell shape and cellular mechanical properties. To investigate these mechanistic links we are further developing high-resolution imaging approaches to characterize cell wall structural dynamics on a nanometer scale during tip growth, and population-level growth dynamics that will help us identify mutants exhibiting perturbed mechanical properties and altered cytoplasmic densities**.**

III. Conclusion

We present an imaging-based approach to quantitatively assess cytoplasmic density and its links to cell cycle progression and cell growth pattern. Mutants blocked for cell cycle progression or inhibition of specific biosynthetic pathways using drug treatments allows us to investigate the regulatory links between growth and biomass accumulation and its potential downstream effects on cell shape and mechanical properties.

References

[1] Lancaster O, *et al.* (2013) Mitotic rounding alters cell geometry to ensure efficient biopolar spindle formation. *Developmental cell* **25,** 270-283.

[2] Neurohr G, *et al.* (2019) Excessive cell growth causes cytoplasm dilution and contributes to senescence. *Cell*.

[3] Munder M, *et al.* (2016) A pH-driven transition of the cytoplasm from a fluid- to a solid-like state promotes entry into dormancy. *Elife* **5.**

[4] Bostan E, *et al.* (2016) Variatinal phase imaging using the transport-of-intensity equation. *IEEE transactions on image processing* **25,** 807-817**.**

Acknowledgements: PDO was supported by a postdoctoral fellowships from the Swiss National Science Foundation under grants P2ELP3_172318 & P400PB_180872.

[1]Department of Cell and Tissue Biology, UCSF, San Francisco, CA 94143, USA. E-mail: pascal.odermatt@ucsf.edu

[2]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA USA.

Mining phage genomes for genetic circuit parts

Jai Padmakumar[1], Jonghyeon Shin[2], and Christopher A. Voigt[3]

***Short Abstract* — Complex genetic circuits can be assembled by composing simpler parts, such as transcriptional repressors. Though many repressors are available, their utility is currently limited by constraints on orthogonality, dynamic range, and toxicity. By exploiting the natural diversity present in phages and "mining" them for repressors, we can assemble a large library of genetic parts that can be used to create complex circuits. Using parts from three phages, we constructed three NOT gates. This suggests that mining phages for parts is a potentially promising approach for expanding the set of quality genetic circuit parts.**

***Keywords* — Part mining, genetic circuits, synthetic biology, cellular programming**

I. Purpose

LIVING organisms possess a wide array of fascinating and useful abilities. Spore-forming bacteria encapsulate themselves nearly indestructible shells allowing them to survive extraordinarily harsh conditions. Nutrients and information flow through massive "superhighways" of underground fungal networks, connecting and supporting the growth of surface plants. To accomplish these feats, cells synthesize external information via intricate regulatory networks, 'computing' the correct decision to make. Exploiting biology's capabilities for our own purposes requires the ability to engineer regulatory control of living cells. This control can be achieved via synthetic cellular computing devices, which function similar to electrical computers, but can be programmed and used to implement *in vivo* logic.

Regulatory logic is commonly achieved in bacteria through the use of constitutive promoters and transcriptional repressors. A classic example of such a system is the *lac* operon, which functions to induce lactose catabolism. Using solely a repressor and an activator, the cell is able to precisely activate expression only when required. In Boolean logic terms, the cell is performing a (lactose *NIMPLY* glucose) operation.

Digital logic provides a neat framework from which to assemble complex computing devices, which electrical engineers have used to design circuits for decades. It uses the formalism provided by Boolean algebra, where a variable has only two values; it can be 1 or 0, or ON and OFF, respectively. Smaller parts called "logic gates", responsible for a single logical operation, are combined to implement sophisticated computations. NOR gates are "universal," meaning all possible computations can be performed using just these gates.

Biologically, a part that can perform a NOR operation can be built using two promoters each regulated by their own repressors. Here, each promoter is a variable, with ON/OFF states defined by whether or not it is being transcribed. NOR gates can be designed from simpler NOT gates, which simply invert the input. NOT gates are easily designed and characterized by using a repressor under control of an inducible promoter and a cognate repressible constitutive promoter controlling expression of a fluorescent output. This function can be characterized by a simple 1-dimensional input-output curve. By characterizing the transfer functions of two NOT gates, we can reliably predict the 2D-transfer function of a NOR gate with a simple model. Our lab has previously developed software to build genetic circuits from characterized NOT gates.[1] Additionally, we have previously constructed a set of NOT gates based on TetR homologs, however these gates often suffer low dynamic range and high toxicity, limiting the complexity of the genetic circuits.[2]

Phage cI repressor has several attributes that make it ideal for logic gate development. cI homologs are well conserved and readily identifiable in other similar phages. Its relatively small size facilitates cloning. cI binds its operator very tightly[3] and preliminary data show it displays high dynamic range, is non-toxic, and has minimal cross-talk due to evolutionary pressure for DNA-binding diversification.

II. Conclusion

By systematically engineering NOT gates from both characterized and uncharacterized phage repressors, we can engineer a large set of high performance NOT gates. 1000s of phages exist in RefSeq database today. Though these repressors do not have known operators, they can be characterized through a combination of bioinformatic and experimental methods. Ultimately, these new parts will be used to construct a large, reprogrammable regulatory circuit.

References

[1] Nielsen, A. A. K. *et al*. Genetic circuit design automation. *Science* **352,** aac7341–aac7341 (2016).

[2] Stanton, B. C. *et al.* Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat Chem Biol* **10,** 99–105 (2013).

[3] Shea, M. A. & Ackers, G. K. The OR control system of bacteriophage lambda. *Journal of Molecular Biology* **181,** 211–230 (1985)

[1]Microbiology program, Massachusetts Institute of Technology, E-mail: jaip@mit.edu

[2]Department of Bioengineering, Massachusetts Institute of Technology

[3]Department of Bioengineering, Massachusetts Institute of Technology E-mail: cavoigt@gmail.com

Rescaling mutant Ras signals by ERK

Michael Pargett[1*], Taryn E. Gillies[1*], Carolyn Teragawa[1], Jillian M. Silva[2], Frank McCormick[2,3], and John G. Albeck[1]

***Short Abstract* — Activating Ras mutations are prominent drivers of cancer, though how they affect signal processing remains uncertain. Using cell lines expressing only one Ras isoform each, including mutants, we analyze signal processing through ERK activity in live cells. Regardless of Ras hyperactivation, ERK activity remains bounded within wild type ranges, and only moderately elevated when unstimulated. The Ras-ERK network drastically attenuates hyperactive signals, yet retains responsiveness to growth factors, through regulation of both network components and the phosphatases that target ERK substrates. These findings imply that the network actively rescales signals to match the input dynamic range, reminiscent of dose-response alignment.**

***Keywords* — Ras, ERK, cell signaling, signal processing, dose-response alignment, feedback, phosphatase, live-cell imaging, single-cell kinetics.**

I. Background

THOUGH activating mutations in Ras are present in ~30% of human tumors, the quantitative effects of these mutations on effector pathway signaling remain uncertain, with activating Ras mutants linked to both increased and decreased ERK activation. As pathway-specific treatment strategies rely on activity estimates at the cellular level, it is increasingly crucial to clarify the specific effects of Ras mutation on downstream pathways.

Originally expected to be vastly hyperactivating, the mutational status of Ras is poorly correlated with average phospho-ERK levels in both tumor cell lines and mouse models [1-2]. Conversion of a wild type Ras gene to a mutant can even reduce average ERK activation [2]. Potential explanations may involve interactions with wild type Ras, additional mutations in cancer, feedback regulation, high cell-to-cell heterogeneity, and/or limited experimental resolution.

II. Approach

We characterize the specific effects of oncogenic Ras mutations on downstream signaling at the single-cell level, using live-cell imaging with an ERK kinase activity sensor in cell lines expressing only one Ras isoform.

A Ras knockout mouse embryonic fibroblast (MEF) line that does not express H-, K-, or N-Ras was previously developed, allowing for isolated exogenous expression of single Ras isoforms [3]. Select Ras isoforms were expressed from an unregulated viral promoter, including wild type H-, K-, and N-Ras, and K-Ras G12C, G12V, G12D, and Q61R.

To comprehensively analyze single-cell Ras/ERK signaling and the status of the whole pathway, we collected a dataset of protein expression, ERK activation (via phospho-ERK level and fraction), and ERK substrate activation (via EKAR3, a genetically encoded ERK activity sensor). Measurements are made in time courses including the absence and presence of stimulation by growth factors. We present analysis of this dataset to identify, for each mutant isoform, statistically relevant signaling differences at the average and single-cell levels, to evaluate the effects and strength of feedback regulation, and to construct a refined model of the pathway-modulated effects of altered Ras function.

We find that oncogenic Ras mutations restrict the range of ERK output, with elevated ERK kinase activity only in the absence of growth factor stimulus, despite excess levels of active Ras. Individual cells expressing mutant Ras proteins are variably responsive to acute growth factor stimulation, but do not exceed the peak magnitude of the wild type.

III. Conclusion

Overall, pathway-level effects including loss of responsiveness, variable negative feedback strength, and ERK substrate-level phosphatase activity serve to rescale changes in ERK activation consistent with the principle of dose-response alignment, thereby preventing excess activity and amplifying changes due to growth factor stimulus. This systematic study reconciles seemingly inconsistent reports within the literature and implies that the initial signaling changes induced by Ras mutations in oncogenesis are inherently subtle.

References

[1] Yeh, JJ, et al. (2009). KRAS/BRAF mutation status and ERK1/2 activation as biomarkers for MEK1/2 inhibitor therapy in colorectal cancer. *Mol Cancer Ther* **8**, 834-843.

[2] Tuveson, DA, et al. (2004). Endogenous oncogenic K-ras(G12D) stimulates proliferation and widespread neoplastic and developmental defects. *Cancer Cell* **5**, 375-387.

[3] Drosten, M, et al. (2010). Genetic analysis of Ras signalling pathways in cell proliferation, migration and survival. *EMBO J* **29**, 1091-1104.

[1] Department of Molecular and Cellular Biology, University of California, Davis, CA. E-mail: mpargett@ucdavis.edu, jgalbeck@ucdavis.edu

[2] UCSF Helen Diller Family Comprehensive Cancer Center, San Francisco, CA.

[3] Frederick National Laboratory for Cancer Research, Frederick, MD, USA.

[*] Equal contributions.

Acknowledgements: This work was funded by NIH grant 1R01GM115650, and DOD grant W81XWH-16-1-0085.

Calcium signals in embryo sidedness and includes Nitric Oxide formation in regulation of normal Heart Loop Morphology

Ashley Pereira, B.S., and Wilfred Denetclaw, Ph.D.

Department of Biology, San Francisco State University, San Francisco, CA 94132

Asymmetric elevation Ca2+ signaling at the left side of the Hensen's node in early embryo development is essential for normal right side heart looping and consequently normal heart development. When calcium levels are elevated on the right side or blocked on the left side, both result in 30% of embryos with situs inversus (reversal of the heart loop). Nitric oxide (NO) formation also depends on elevation of calcium and can promote cardiomyocyte proliferation. However the role of NO in regulation of heart looping is unknown. Therefore, we hypothesize NO involvement in cardiac looping may be by its ability to regulate cardiomyocyte proliferation and under absence of NO may cause a looping defect due to absence of sufficient cardiogenic cells.To investigate, acrylic beads soaked in L-NAME, a competitive inhibitor of nitric oxide synthase(NOS), was placed to the right side of Hensen's Node to block NO signaling but did not affect Ca2+ change. Our findings show 90% abnormal heart tube formation with situs ambiguous, a range of morphological heart looping defects. These results suggests that the calcium elevation on the right side functions in the activation of NOS for NO production and signaling, possibly by its canonical pathway, to regulate in the normal heart loop. NO signaling in early embryo development is not well understood and it's role in cardiomyocyte proliferation may also need to be studied to fully understand cardiogenesis..NIH MBRS-RISE: R25-GM0592
NSF STC CCC: 154829798

Classification using Expectation Reflection

Danh-Tai Hoang[1], Junghyo Jo[2], and Vipul Periwal[1]

***Short Abstract* — Data classification is a fundamental problem in biology. We recently developed a data-driven approach, Expectation Reflection, for network inference. The method outperformed previous methods in predicting interactions underlying observations of stochastic processes. Using hidden variables, we extend the method to solve unsupervised classification problems. Applying our method to the MNIST handwritten digit dataset, we demonstrate that we can successfully classify this data into 60 clusters. Interestingly, the numbers of clusters for each digit were approximately the same.**

***Keywords* — Classification, Network inference, Expectation Reflection, hidden variables.**

I. Purpose

Classification of complex data into different clusters is a critical topic, not only in quantitative biology but more generally in data science. We recently developed a new approach to data-driven inference, Expectation Reflection (ER), and demonstrated that ER outperforms other existing methods in recovering interaction weighs between variables from sequential data, especially in the limit of small sample sizes and partially observed systems [1, 2].
By adding hidden variables and considering interactions from hidden variables to observed variables, we extend the ER method to unsupervised classification.

II. Method

We added categorical hidden variables to the system and considered only interactions from the hidden variables to observed variables. Our algorithm contains the following steps: (i) Assign the state of hidden variables as random; (ii) Apply ER [1, 2] to infer interaction strengths from hidden variables to observed variables; (iii) Update the state of hidden variables to the value that maximizes the likelihood of the observations; and (iv) Repeat steps (ii) and (iii) until the discrepancy of the observations and the corresponding model expectations saturates. The number of hidden variables (i.e., number of clusters) can be estimated from the minima of the discrepancy of the entire system.

Acknowledgements: This work was supported by the Intramural Research Program of the National Institutes of Health, NIDDK (D.-T.H., V.P.).

[1]Laboratory of Biological Modeling, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. Email: danh-tai.hoang@nih.gov, vipulp@mail.nih.gov

[2]School of Computational Sciences, Korea Institute for Advanced Study, Seoul 02455, Korea. Email: jojunghyo@kias.re.kr

III. Results

As a demonstration to a problem with interpretable results, we applied ER to classify the 60,000 images of MNIST handwritten data without knowing the labels (i.e., unsupervised learning). We found that the discrepancy of entire system reached a minimum at about 60 hidden variables. This suggested that we can optimally classify the data into 60 clusters (Figure 1B). The mean images for each cluster are shown in Figure 1C. We then applied this approach to classify single cell expression data.

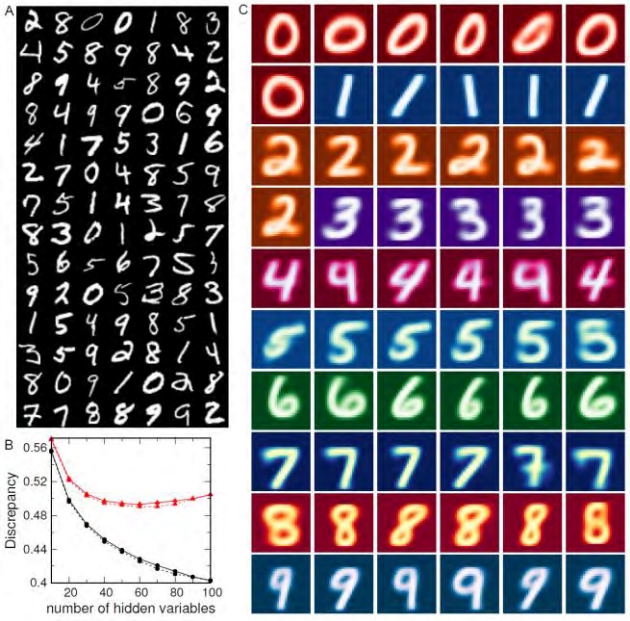


Figure 1. Classification of handwritten digit images. (A) MNIST image samples were randomly selected to demonstrate. (B) Discrepancies are shown as a function of the number of hidden variables, when using 60,000 samples (solid lines) and 20,000 samples (dashed lines). (C) Mean values of each pixel within the same cluster were obtained from our classification.

IV. Conclusion

We extended ER to unsupervised classification problems. We demonstrated the performance of this method by successfully classifying the MNIST data of handwritten digits, and then applied it to single cell datasets.

References

[1] D. T. Hoang, J. Song, V. Periwal, and J. Jo (2019), Network inference in stochastic systems from neurons to currencies: Improved performance at small sample size, Physical Review E, **99**, 023311.

[2] D. T. Hoang, J. Jo, and V. Periwal (2019), Data-driven inference of hidden nodes in networks, arXiv:1901.04122.

Towards a Quantitative Understanding of Spontaneous Mitotic Waves

Owen Puls[1], and Qiong Yang[1,2]

***Abstract*— At the foundation of complexity in biology is the process by which cells proliferate. Regulating this process facilitates the longevity of multicellular organisms by enforcing a regular, clock-like timing of mitotic events: we can consider any cell as a single oscillator. Alternatively, when a collection of these oscillators couple via diffusion, they synchronize. However, at the embryo scale, a diffusion-mediated signal itself is too slow to synchronize mitosis. Instead, waves constitute a tenable spatial coordination mechanism. This work probes mitotic waves: offering preliminary results in adapting a minimal excitable system to identity their properties and basic principles.**

***Index Terms*— self-organization, coupled oscillators**

I. Introduction

Across many years, various models–and adaptations thereof–were developed to capture the mechanism behind the mitotic clock [1]. In short, various checkpoints controlled by the activity of certain compounds direct the cell through a series of steps which define one mitotic cycle. The regulation of these processes facilitates the longevity of multicellular organisms by enforcing a regular, clock-like timing of mitotic events [1].

In particular, the activity of cyclinB-Cdk1 (cyclin-dependent kinase) complexes controls the mitotic state of a cell by promoting the activity of anaphase-promoting compound–APC–which itself inhibits the complex and promotes mitotic exit [2]. These complexes themselves regulate, and are regulated by, Cdc25 and Wee1 in the form of double positive and negative feedback loops, respectively [2]. Inside any singular cell, therefore, we can think of the system as constituting a single oscillator.

In contrast, when a collection of these oscillators couple via diffusion, they synchronize: a phenomenon realized experimentally [1]. For example, in various systems–e.g. *Drosophila* and *Xenopus*–early embryogenesis is marked by a series of synchronous cell divisions across the length of the embryo [3], [4]. Some of these eukaryotic cells can reach up to 600 μm in radius. As such, a hypothetical diffusion-mediated signal (with $D_p = 10\mu$m/s) originating from the center would take hours to reach the edge of the cell [3]. However, given the synchrony observed experimentally, such signal must manage the distance in minutes, not hours [3]. This known asks the question: how does the cell accomplish this feat?

[1]Department of Physics, University of Michigan. Email: owpuls@umich.edu

[2]Department of Biophysics, University of Michigan. Email: qiongy@umich.edu

To answer this question, recent work demonstrates a tenable spatial coordination mechanism at the relevant length scale takes the form of a chemical wave [3], [4]. Multiple models describe one-dimensional mitotic waves, displaying diverse forms and characteristics [1], [2], [3], [4].

II. Results

Using an *Xenopus* extract system, we reconstitute mitotic waves in one-dimensional Teflon-coated tubes, in two different contexts [3], [5]. First, spontaneous waves are visualized naturally as in Ref [3]. Second, inspired by work studying apoptotic waves, we utilize CSF extract (metaphase arrest) as a source to produce directed waves [6].

Moreover, we take advantage of both extracts containing reconstituted nuclei and those without (chemical). Via nuclei markers, we visualize waves of nuclear-envelope formation and breakdown; conversely, from various fluorescent protein signals, we can directly visualize waves of chemical mitotic activity. We believe this is the first realization of the latter.

In both instances, we quantify the wave speed–the only meaningful wave property in one spatial dimension–and find a slowing down over time, similar to period lengthening previously reported. Further, comparing the two, we find a significant difference in their speeds, possibly due to the nuclear-translocation of clock constituents.

III. Conclusions

In total, we report experimental realizations of both nuclear and chemical mitotic waves *in vitro*. Quantification of their respective waves speeds elucidates significant differences between the classes. Due to their relative ubiquity, this apparent distinction may shed light on mitotic waves in diverse biological sysytems moving forward.

References

[1] B. Novak and J. Tyson, "Numerical analysis of a comprehensive model of m-phase control in *Xenopus* oocyte extracts and intact embryos," *J Cell Sci*, vol. 106, pp. 1153–1168, 1993.

[2] Q. Yang and J. Ferrell, Jr, "The cdk1-apc/c cell cycle oscillator circuit functions as a time-delayed, ultrasensitive switch," *Nat Cell Bio*, vol. 15(5), pp. 519–525, 2013.

[3] J. Chang and J. Ferrell, Jr, "Mitotic trigger waves and the spatial coordination of the *Xenopus* cell cycle," *Nature*, vol. 500, pp. 603–607, 2013.

[4] M. Vergassol, V. Deneke, and S. Di Talia, "Mitotic waves in the early embryogenesis of *Drosophila*: Bistability traded for speed," *PNAS*, vol. 115(10), pp. E2165–E2174, 2018.

[5] Y. Guan *et al.*, "A robust and tunable mitotic oscillator in artificial cells," *eLife*, vol. 7, p. e33549, 2018.

[6] X. Cheng and J. Ferrell, Jr, "Apoptosis propagates through the cytoplasm as trigger waves," *Science*, vol. 361, pp. 607–612, 2019.

Measuring the Influence of *cis*-Acting Changes on the Transcriptional Response to Infection in *Drosophila*

Bryan Ramirez-Corona[‡,1] and Zeba Wunerlich[‡,2]

Insects respond to infections by regulating gene expression through several signaling pathways, this response is highly variable between individuals. Past efforts to elucidate this variability focused on protein-coding regions. However, ample variation exists in non-coding DNA sequences of *Drosophila melanogaster*, and in other contexts, this variation has been shown to drive phenotypic differences between individuals. We measure the contribution of *cis*-regulatory sequences to transcriptional variation, by performing tissue-specific allele-specific expression analysis on two inbred strains and their F1 hybrids. By comparing expression patterns pre- and post-infection, we can characterize genetic causes of tissue-specific expression variation in two distinct biological states.

***Keywords* — Gene regulation, transcription, variation, drosophila, immune.**

In eukaryotes, gene expression is controlled by the interactions of *trans*-acting factors, such as transcription factors, and *cis*-regulatory elements (CREs), such as enhancers. Variation in either protein-coding regions or CREs can change expression patterns. Currently, there is a much deeper understanding of how changes in protein-coding regions contribute to expression variation than in *cis*-regulatory elements. There is a need to quantify the contributions of CREs in highly variable systems.

The *Drosophila* immune system shows genetically encoded variation in pathogen resistance and transcriptional response to infection [1,2,3], due to constant exposure to continuously evolving pathogens [4]. In whole animals, there are ample contributions of *cis*-regulatory sequences both at rest and in response to infection [5,6,7]. Since *cis*-regulatory sequences tend to act in a tissues-specific manner, the whole-body analysis may miss effects that are localized. Therefore, there is a need to quantify allele-specific expression in specific tissues. Here we focus on the fat body, the primary immune-responsive organ in *Drosophila.*

To measure these contributions, we generated F1 hybrids by crossing two highly inbred, genetically divergent *D. melanogaster* lines from the *Drosophila* Synthetic Population Resource. We infected flies with Gram-negative bacteria *Serratia marcescens*, a naturally occurring pathogen of *Drosophila.* Three hours after inoculation, we measured gene expression in dissected fat bodies of both the F1 hybrids and parental strains. By comparing allelic ratios in the F1 hybrid to the parental strains, we can identify genes whose differential expression is caused by changes in *cis* or *trans.*

By measuring allele-specific expression in the fat body we will, for the first time, quantify *cis* and *trans* effects on the immune response in a high-resolution, tissue-specific manner. We expect to find many genes with changes in *cis*-regulatory regions that account for changes in transcriptional response to immune stimulation, due to the evolutionary pressures on the immune system to diversify [4] coupled with the capacity of changes in *cis* to produce large expression differences in a localized manner [8]. Preliminary data supports this prediction. This data set will allow us to answer questions regarding the pattern of gene expression divergence, both at the basal state and after immune stimulation.

This data will open future avenues of investigation, since potential regulatory regions of genes found to carry changes in *cis* can be identified using DNA accessibility assays [9]. Identifying differentially expressed genes due to changes in *cis* along with the corresponding regulatory regions will allows us to identify the specific sequence changes that affect gene expression and generate models of the connection between immune enhancer sequence and function.

References

[1] Sackton, T. B., Lazzaro, B. P., & Clark, A. G. (2010). Genotype and Gene Expression Associations with Immune Function in Drosophila, *6*(1).

[2] Lazzaro, B. P., Sceurman, B. K., Clark, A. G., & Sciences, B. (2004). Genetic Basis of Natural Variation in D . melanogaster. *Science*, *303*(March), 1873–1877.

[3] Lazzaro, B. P., Sackton, T. B., & Clark, A. G. (2006). Genetic Variation in Drosophila melanogaster Resistance to Infection :, *1554*(November), 1539–1554.

[4] Lazzaro, B. P. (2008). Natural selection on the Drosophila antimicrobial immune system, 284–289. http://doi.org/10.1016/j.mib.2008.05.001

[5] Wittkopp, P. J., Haerum, B. K., & Clark, A. G. (2004). Evolutionary changes in cis and trans gene regulation. *Nature Publishing Group*, *430*(1), 85–88.

[6] Osada, N., Miyagi, R., & Takahashi, A. (2017). Cis- and trans-regulatory effects on gene expression in a natural population of Drosophila melanogaster. Genetics, Vol. 206, 2139–2148

[7] Juneja, P., Quinn, A., & Jiggins, F. M. (2016). Latitudinal clines in gene expression and cis -regulatory element variation in Drosophila melanogaster. *BMC Genomics*, 1–11.

[8] Kvon, E. Z., Kamneva, O. K., Melo, S., Dickel, D. E., Melo, S., Barozzi, I, Pennacchio, L. A. (2016). Progressive Loss of Function in a Limb Enhancer during Snake Evolution, 633–642.

[9] Daugherty, A. C., Yeo, R. W., Buenrostro, J. D., Greenleaf, W. J., Kundaje, A., & Brunet, A. (2017). Chromatin accessibility dynamics reveal novel functional enhancers in C . elegans. *Published by Cold Spring Harbor Laboratory Press*, (27), 2096–2107.

‡Developmental and Cell Biology, University of California, Irvine
[1]E-mail: baramir2@uci.edu
[2]E-mail: zeba@uci.edu

Inferring a Lot at Once from the Same Cell

GW McElfresh[1,*], Brian Drawert[2,*], and Christian J. Ray[1,3,*]

***Short Abstract* — High-dimensional single-cell resolution datasets represent the most quantitative information about the state of the same cell that has ever been experimentally attainable by orders of magnitude. Using a complexity-aware modeling approach, we have developed an *E. coli* cell cycle simulator that exploits such large datasets to model the stochastic trajectories of a regulatory network of interest coupled to a hybrid statistical/stochastic model of the cell cycle that accounts for metabolic costs of the network, gene copy number, cell volume, and other factors. We implemented a detailed model of the PhoBR two-component system that responds to phosphate starvation, synchronized independent cell cycle runs, and quantified the relationships between molecular fluctuation of each species or other model quantity, such as growth rate and volume. Our results reveal several surprising relationships between the timescales of effects.**

***Keywords* — complexity-aware modeling, high-dimensional data, cell cycle, stress response, two-component system**

I. Background

High-dimensional single-cell datasets are being generated in transcriptomics, time-lapse microscopy, and, increasingly, proteomics, at an increasing rate. They have contributed to the revitalization of the question of cell type identity and stand to aid the functionalization of polygenic genetic traits. Among the insights that can be gained, the quantity of data already available or feasibly attainable are allowing new approaches to whole-cell predictive simulation [1]. While comprehensive mechanistic whole-cell models have many challenges, the development of complexity-aware modeling provides an opportunity to quantitatively analyze long trajectories at single-cell resolution, trajectories from many cells, and cellular lineages [2-4].

We are developing a data-driven bacterial cell cycle algorithm that couples the outputs of a small subnetwork of interest to global cellular physiology in high detail, including chromosome replication, gene copy number, sequence-specific timing of transcription, cell volume, and any other parameters of cellular physiology relevant to the network of interest. The framework uses the adder principle [5] to determine cell size control.

The phosphate starvation stress response in *E. coli* is the first network implemented in the framework, chosen because of its extensive characterization on mechanistic and physiological levels [6, 7]. At the sensing step is a two-component system, PhoBR, that alters a regulon of approximately 50 genes. Our network model uses quantitative stochastic simulation with precise parameters, realistic transcriptional and translational delays, and an energy budget that feeds into the cell cycle model. All regulated genes in the phosphate starvation response are included at energetic costs proportional to their size.

II. Results

A. Multiple Timescales of Response

By capturing the system in its physiological context and enforcing local rates to be correct, we observed distinct timescales emerging from simulations, from molecular conformational fluctuation to multiple cell cycles.

B. Distinct Response Phases During the Cell Cycle

We synchronized the cell cycles of independent simulation runs to determine the effect of phosphate response fluctuations on the cell cycle and the response itself, including subspecies that are not experimentally quantifiable at single-cell resolution. Using sensitivity analysis and principal component analysis, we identified that processes cluster into discrete fluctuation rates, from fast fluctuations in the effect of active regulator to slower changes in the effect of cell volume on the response. Notably, the effect of cell volume evolves with the cell cycle, with stronger absolute effects shortly before and after cell division.

III. Conclusion

Physiological context is essential for predictive network simulations. The hybrid approach reveals fundamental new insights and relationships in gene regulation.

References

1. Karr JR, *et al.* 2012. Cell. 150(2):389-401.
2. Gomez-Schiavon M, El-Samad H. 2018. Curr Opin Microbiol. 45:47-52.
3. Bandyopadhyay A, Wang H, Ray JCJ. 2018. PLoS Comput Biol. 14(8):e1006380.
4. McElfresh GW, Ray JCJ. Intergenerational Cellular Signal Transfer and Erasure. The Energetics of Computing in Life and Machines: SFI Press 2019.
5. Jun S, *et al.* 2018. Rep Prog Phys. 81(5):056601.
6. Gao R, Stock A. 2013. Proc Natl Acad Sci U S A. 110(2):672-7.
7. Marzan LW, Shimizu K. 2011. Microb Cell Fact. 10:39.

Acknowledgements: This work was funded by NIH grant P20GM103638.
[1]Center for Computational Biology, University of Kansas, Lawrence, KS
[2]Computer Science, University of North Carolina, Asheville, NC
[3]Molecular Biosciences, University of Kansas, Lawrence, KS
*Equal contributions from all co-authors.
Contact: gwmcelfresh@ku.edu, bdrawert@unca.edu, jjray@ku.edu

Mitochondrial Localization Within *Stentor*

Nicole Rodrigues[1], Rebecca McGillivary[1], and Wallace Marshall[1]

***Short Abstract* — *Stentor coeruleus* is a eukaryotic, unicellular ciliate. It can completely regenerate once it is cut in half. There is almost nothing known about the membrane-bound organelles within *Stentor* except for the micronuclei and macronucleus. To examine organelle distribution within *Stentor* during regeneration, the organelles first have to be located and characterized. To begin, I fluorescently labeled mitochondria in *Stentor* with a MitoTracker dye to study the distribution of mitochondria. I observed a network of puncta throughout the Stentor between 2-4 h after using the MitoTracker. The puncta are within a 1.3 µm to 3.5 µm size range.**

I. Introduction

Stentor coeruleus is a unicellular eukaryote. It is a 1 mm long, trumpet-shaped ciliate that has the ability to completely regenerate once it is cut in half. *Stentor* has complex and unique distinguishing structures at both its posterior end and anterior end that can easily be seen under a light microscope [1], such as the oral apparatus (the organelle that enables *Stentor* to engulf food), the cortical rows (which give *Stentor* its trumpet-shape and are located at the cortex), the micronuclei and a long macronucleus. The micronuclei and macronucleus are the only membrane-bound organelles where the localization is known. The only information known about the mitochondria in *Stentor* is from electron microscopy (EM) images which give the mitochondria size (~1 µm) and show the mitochondria near the cortex [2]. I characterized mitochondria localization in *Stentor*.

II. Methods and Results

To study mitochondrial localization in *Stentor*, I used MitoTracker Green dye to label the mitochondrial membrane, and fluorescent dextran was used as a control. *Stentor* can ingest the dyes that it is being incubated with, and the dye will end up in the food vacuoles. Dextran was used to find *Stentor*'s food vacuole size range (3.5 µm to 50 µm). Thus, I was able to rule out the MitoTracker puncta observed within the food vacuole size range. 30 min after incubating the *Stentor* with the MitoTracker and dextran, puncta were observed (1.3 µm to 22.5 µm size range). A network of puncta is visible throughout the *Stentor* between 2-4 h after using the MitoTracker— the puncta are located at the oral apparatus and the cortical rows. The puncta are enriched at the cortex and they are within a 1.3 µm to 3.5 µm size range. This data is consistent with the size range of the electron microscopy images of the mitochondria in *Stentor* [2].

III. Future Directions

There are many unknowns surrounding mitochondria. It is currently unknown if changes to the mitochondrial distribution are involved in eukaryotic cell response to perturbations [3]. As *Stentor* can regenerate after being cut in half, measuring the mitochondrial distributional from this perturbation in *Stentor* could help address those unknowns.

Currently, as the mitochondria can be visualized within *Stentor*, experiments are underway to investigate if the mitochondrial distribution changes during regeneration. The videos of live *Stentor* that are regenerating can be computationally analyzed to measure the network's distributional changes via looking at features such as size and position to compare the mitochondrial distribution prior, during and after regeneration [4]. I will look at how computational models such as bacterial growth simulations could mimic how the mitochondria regenerate their numbers when *Stentor* is cut in half. All this mitochondrial data can help make *Stentor coeruleus* an amazing model to use to understand how organelles regenerate and behave within a morphologically complex single-celled organism with amazing regenerative capabilities.

References

[1] Tartar V (1961) *The Biology of Stentor.* Oxford, Pergamon.
[2] Huang B & Pitelka DR (1973) THE CONTRACTILE PROCESS IN THE CILIATE, STENTOR COERULEUS. *Journal of Cell Biology*.
[3] Hoitzing H, et. al. (2015) What is the function of mitochondrial networks? A theoretical assessment of hypotheses and proposal for future research. *Bioessays* **37**, 687-700.
[4] Rafelski SM (2013) Mitochondrial network morphology: building an integrative, geometrical view. *BMC biology*, 11, 71.

Acknowledgements: This work was funded by NSF grant DBI-1548297.

[1]Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143, USA. E-mail: Nicole.Rodrigues@ucsf.edu, Rebecca.McGillivary@ucsf.edu, Wallace.Marshall@ucsf.edu

Computation Using Patterning of Bacterial Colonies

Luca Rosa[1], Paul Grant[3], Alex Fedorec[1], Andrew Phillips[3], and Chris P Barnes[1,2]

***Abstract*— Whole-cell bacterial biosensors are living organisms that implement detection of a specific analyte and produce an easily detectable and quantifiable response. Some example are biosensors detecting heavy metals such as arsenic, or health related metabolites such as lactate. Diseases are complex with many sharing the same symptoms and biomarkers. Therefore there is a need of biosensors capable of combining and processing multiple inputs for greater discriminatory power.**

I. Extended abstract

Biosensors are devices that consist of a living organism that implement a system able to detect a specific analyte and then produce and easily detectable and quantifiable response [1]. The development of biosensors is currently an important focus of synthetic biology. Different whole-cell biosensors have been developed where a system detecting an analyte is integrated into a living organism [2], [3]. The limitation of this approach is that they usually sense a single metabolite. Although genetic logic gates can be engineered within single-cells, a new strain needs to be engineered if a new function is required. There is therefore the need for a platform where different biosensors can be combined to perform multiplexed computation with different inputs.

The aim of the project is to build a platform of engineered bacterial colonies capable of performing computation through spatial pattern formation. The architecture consists of colonies arranged in a grid structure where each circle represents a node of the system. Two types of strains are implemented: the input and the output strains. The former integrate the biosensor system and produces N-Acyl homoserine lactone (AHL), an intercellular signalling molecule, which diffuses in the agar [4]. The latter expresses a reporter protein in response to AHL molecules. The computation is performed spatially, where the final pattern is a direct consequence of the metabolites present in the sample and of the initial arrangement of the input nodes.

Here we a demonstrate a proof-of-principle system based on an arabinose induced input strain and an output strain that responds monotonically to AHL. Using a computational model, we also demonstrate that integrating different switching mechanisms in the output strain results in different spatial logic functions. In the future, we want to design geometries that can perform multiple signal integration and parallel logic functions.

[1]University College London - Cell and Developmental Biology
[2]University College Genetics Institute (UGI)
[3]Computational Science Laboratory, Microsoft Research, Cambridge, UK

ACKNOWLEDGMENT

This project was supported by the LIDo PhD consortium and funded by Microsoft Research and BBSRC

References

[1] Institute of Medicine (US) Forum on Microbial Threats, "The Science and Applications of Synthetic and Systems Biology: Workshop Summary," 2011.

[2] E. Diesel, M. Schreiber, and J. R. van der Meer, "Development of bacteria-based bioassays for arsenic detection in natural waters," *Analytical and Bioanalytical Chemistry*, vol. 394, no. 3, pp. 687–693, Apr. 2009.

[3] L. Goers, C. Ainsworth, C. H. Goey, C. Kontoravdi, P. S. Freemont, and K. M. Polizzi, "Whole-cell Escherichia coli lactate biosensor for monitoring mammalian cell cultures during biopharmaceutical production." *Biotechnology and Bioengineering*, vol. 114, no. 6, pp. 1290–1300, Jun. 2017.

[4] C. M. Waters and B. L. Bassler, *Quorum sensing: Cell-to-cell communication in bacteria*. Princeton University, Princeton, United States: Annual Reviews, Dec. 2005, vol. 21.

A Time-Dependent Evolutionary Strategy to Discover Generalist Genotypes

Vedant Sachdeva[1], Kabir Husain[2], Shenshen Wang*, Arvind Murugan*[2]

***Short Abstract* —Natural environments can present diverse fitness pressures, but some genotypes remain fit across a wide range of challenges. Such `generalists' can be hard to evolve because there may be entropic or absolute fitness costs relative to specialists. On the other hand, it is critical to evolve such generalists, for example, in the search for broadly neutralizing antibodies during affinity maturation. Here, we study the conditions under which time-dependent evolutionary protocols can evolve generalists even when static protocols fail. We find that cycling environments on timescales tuned to match fixation times can reliably evolve generalists when either deleterious selection is too adverse or when fitness landscapes are too rugged due to epistasis for static protocols to succeed. We discuss `chirp' protocols that can more reliably produce larger generalist populations. Our work reveals regimes in which time-dependent `seascapes' can dynamically enlarge the effective attractor size around genotypes that are hard to find in any static protocol.**

***Keywords* — changing environments, immunology models, population genetics models**

I. Purpose

TIME-VARYING selection pressures as a result of fluctuating environments drive evolution out of equilibrium, resulting in the emergence of non-equilibrium stationary states[1]. These non-equilibrium stationary states describe behaviors not achievable in equilibrium evolutionary conditions, such as the emergence of bet hedging, phenotype switching, and generalist phenotype strategies[2] These strategies are the effort of an evolving system to optimize a tradeoff between being well-specialized for one environment against being well-specialized to another environment. Here, we take advantage of this phenomenon in order to drive evolutionary systems towards generalist genotypes, in spite of the fact that the generalist genotypes are not accessible in equilibrium conditions.

Evidence of the ability of time-varying selection pressures to bring about generalist genotypes has been observed in the affinity maturation of B-cells against strains of HIV[3]. Computational studies of the affinity maturation process demonstrated that B-cells are more likely to evolve the ability to produce bnAbs when exposed to multiple strains of HIV sequentially in the affinity maturation process rather than exposed to multiple strains of HIV simultaneously. This demonstrates that idea that time-varying environments can drive the discovery of generalizing strategies in cases where static environments fail.

In this work, we will demonstrate that environmental cycling at frequencies near the fixation timescale of a single mutation can drive evolution towards discovering generalist genotypes, even in conditions when no static set of fitness pressures can stabilize populations at that genotype. This effect is demonstrated to be relevant to both rugged and flat fitness landscapes, suggesting a general framework to evolve robust populations. We further demonstrate that this result can be thought of as a tradeoff between two distinct phenomena: discovering the generalist genotype and staying at the generalist genotype. We then demonstrate another kinetic strategy that can break the tradeoff between these two phenomena and provide even better results.

II. Methods and Results

We utilize Yule processes and Moran models to model populations evolving under periodic environmental conditions. The impact of fluctuating environmental conditions is shown to be important when the generalist genotype is disfavored compared to specialist genotypes. By computing both the probability of an individual with a generalist genotype emerging and the probability of the population remaining at the generalist genotype as a function of the environmental period. The tradeoff between the two quantities is optimized using environmental periods close to the length of time needed for the fixation of a single mutation. In addition, environmental strategies that slowly decrease the period are shown to provide an even greater advantage.

We propose that environmental fluctuations enable the population to take advantage of stochasticity during the evolutionary process, enabling it to explore more of the possible genotypes. This suggests that environmental fluctuations play a critical role in developing populations that learn common features.

References

[1] Mustonen, Ville, Lassig, Michael. (2009) From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. *Trends in Genetics*. **25**, 111-119.

[2] Mayer, Andreas, Mora, Thierry, Rivoire, Olivier, Walczak, Aleksandra. (2017) Transitions in optimal adaptive strategies for populations in fluctuating environments. *Phys. Rev. E.* **96,** 032412.

[3] Wang, Shenshen, Mata-Fink, Jordi, et al. (2015) Manipulating the Selection Forces during Affinity Maturation to Generate Cross-Reactive HIV Antibodies. *Cell.* **160,** 785-797.

[1]Graduate Program in Biophysical Sciences, University of Chicago E-mail: sachdved@uchicago.edu

[2]James Franck Institute, University of Chicago E-mail: amurugan@uchicago.edu

Heterogeneity in efflux pump expression in different growth contexts

Nadia Sampaio[1,2], Imane El Meouche[1,2], Jean-Baptiste Lugagne[1,2] and Mary J. Dunlop[1,2]

Antimicrobial resistance is a major public health concern. Recently, our group demonstrated that transient expression of the efflux pump AcrAB-TolC is inversely correlated with expression of the DNA repair protein MutS, thus predisposing cells to acquire mutations that can potentially lead to high-level resistance. This relationship was observed during exponential cell growth in a nutrient-rich environment, which is rarely encountered outside the laboratory. We hypothesize that different growth contexts might amplify or weaken the inverse relationship between these genes. To address this question, we are investigating this relationship at single cell level during growth using experiments that test the impact of the media and growth phase.

***Keywords* — Antibiotic resistance, phenotypic heterogeneity, efflux pumps, mutation rates.**

I. Project Description

Bacteria leverage a variety of mechanisms to survive antibiotic treatment. High-level resistance is primarily achieved through *de novo* mutations or horizontal gene transfer[1]. However, cells can temporarily withstand drug treatment via mechanisms such as cell dormancy, spore formation or efflux pump overexpression[2–4]. Recently, our group demonstrated that one of such mechanisms for transient tolerance can facilitate the acquisition of permanent mutations[5]. *Escherichia coli* cells overexpressing the efflux pump AcrAB-TolC have low levels of the mismatch repair protein MutS and increased mutation rates. This striking relationship was observed in cells growing exponentially with virtually unrestricted access to nutrients. However, conditions that sustain constant bacterial growth are rarely found outside of the laboratory setting. During host infections, and in most natural environments, bacteria have to face fluctuating nutrient availability until favorable growth conditions are restored. Consequently, alternations between periods of exponential growth and starvation are also likely frequent. In this scenario, the cost-benefit relationship between AcrAB expression and elevation in mutation rates might change. Interestingly, AcrAB shows an inverse correlation with growth rate in nutrient-limited media[6] and MutS is downregulated with entry in stationary phase[7]. Here, we hypothesize that different growth contexts might amplify or weaken the inverse relationship between AcrAB and MutS. We will present results characterizing the relationship between the expression profile of these two genes across different growth media and phases of growth. To do so, *E. coli* cells containing a plasmid to report expression simultaneously from the *acrAB* and *mutS* promoters (P_{acrAB}-*rfp* + P_{mutS}-*yfp*) are cultured in increasingly rich media. Using fluorescence microscopy, we quantify cell-to-cell variation in the expression of both reporters and identify how this inverse relationship is altered in specific stages of growth. Finally, to elucidate potential regulatory links between AcrAB and MutS, we follow expression dynamics at the single cell-level using a "mother machine" microfluidic device and apply cross-correlation analysis. This work will shed light on potential mechanisms that cells can leverage to fine tune costly antibiotic tolerance mechanisms and potential acquisition of resistance.

References

1. Alekshun, M. N. & Levy, S. B. Molecular Mechanisms of Antibacterial Multidrug Resistance. *Cell* **128,** 1037–1050 (2007).
2. Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L. & Leibler, S. Bacterial persistence as a phenotypic switch. *Science (80-.).* **305,** 1622–1625 (2004).
3. Pu, Y. *et al.* Enhanced Efflux Activity Facilitates Drug Tolerance in Dormant Bacterial Cells. *Mol. Cell* **62,** 284–294 (2016).
4. El Meouche, I., Siu, Y. & Dunlop, M. J. Stochastic expression of a multiple antibiotic resistance activator confers transient resistance in single cells. *Sci. Rep.* **6,** 1–9 (2016).
5. El Meouche, I. & Dunlop, M. J. Heterogeneity in efflux pump expression predisposes antibiotic-resistant cells to mutation. *Science* **362,** 686–690 (2018).
6. Rand, J. D., Danby, S. G., Greenway, D. L. A. & England, R. R. Increased expression of the multidrug efflux genes acrAB occurs during slow growth of Escherichia coli. *FEMS Microbiol. Lett.* **207,** 91–95 (2002).
7. Feng, G., Tsui, H. C. T. & Winkler, M. E. Depletion of the cellular amounts of the MutS and MutH methyl-directed mismatch repair proteins in stationary-phase Escherichia coli K-12 cells. *J. Bacteriol.* **178,** 2388–2396 (1996).

Acknowledgements: This work was supported by the National Institutes of Health grant 1R01AI102922.

[1] Department of Biomedical Engineering, Boston University
[2] Biological Design Center, Boston University
Presenting author contact info: nsampaio@bu.edu

Control Structures of a Cancer Drug Resistance Model

James M. Greene[1], Cynthia Sanchez-Tapia[1], Jana L. Gevertz[2] and Eduardo D. Sontag[3]

***Abstract*— Drug resistance is one of the main factors limiting the success of chemotherapy in cancer treatment. Its mechanisms are classically understood as conferred to the cell by random genetic mutations, from which clonal expansion occurs via Darwinian evolution. However, the more recent experimental discovery of epigenetic and phenotype plasticity suggests that chemotherapy can produce drug-resistant clones. In this work, we seek a treatment protocol which maximizes the time to reach a critical tumor size. Utilizing differential-geometric techniques, the control structure is characterized as a concatenation of bang-bang and path-constrained arcs.**

I. Description

Drug resistance is a major factor limiting the success of cancer chemotherapy. In general, resistance can be classified as either *intrinsic* or *acquired* [1]. Intrinsic resistance refers to the case when resistant subpopulations exist prior to treatment, and are subsequently selected by the drug in a Darwinian manner. Conversely, acquired resistance refers to resistant clones that are generated during the course of therapy. Resistance mechanisms may be *randomly acquired* due to genetic mutations and/or epigenetic phenotype-switching [2], or may be *induced* by the presence of the drug itself [3], [4], [5], [2]. For example, Pisco and colleagues measured the relative contribution of resistance selection (pre-existing and randomly acquired) versus drug-induced resistance in HL60 leukemic cells [2]. Both pre-existing and randomly acquired resistance have been well studied mathematically. However, *drug-induced resistance* is a more recently discovered experimental phenomenon and lacks substantial theoretical analysis.

In a previous work we have developed and analyzed a mathematical framework to distinguish between the forms of resistance mentioned above. Given a fixed rate of resistance induction, we explored the impact of this parameter on treatment outcomes [6], as well as the resulting consequences for how to schedule chemotherapy [7], [8]. In these works we assumed that the cancer population is composed of two types of cells: sensitive (S) and resistant (R). For simplicity, the drug is taken as completely ineffective against the resistant population, while the log-kill hypothesis is assumed for the sensitive cells. Complete resistance is of course unrealistic, but can serve as a reasonable approximation, especially when toxicity constraints are considered. The effect of treatment is considered as a control agent $u(t)$, which we assume is a locally bounded Lebesgue measurable function taking values in $\mathbb{R}_+$. Here $u(t)$ is directly related to the applied drug dosage.

We utilize both the Pontryagin Maximum Principle and differential-geometric techniques to characterize solutions that maximize the time until treatment failure. The necessary conditions then imply that the optimal control can be synthesized as a combination of bang-bang and path-constrained arcs. Numerical results that support the computed control structure are provided. We also investigate the dependence of the control structure and treatment efficacy as a function of both the chemotherapeutic cytotoxicity (a classical measure of the effectiveness of treatment) and the rate at which resistance is induced by the drug. Our results suggest that the latter may significantly alter the outcome of treatment, and may in fact be more important than drug toxicity in certain parameter ranges. Hence, the propensity of a treatment to promote resistance is clinically significant, demonstrating the need for further experimental and mathematical research.

This work was funded partially by grants DARPA FA8650-18-1-7800, NSF 1817936, and AFOSR FA9550-14-1-0060.

[1] Department of Mathematics and CQB, Rutgers University, Piscataway, New Jersey 08854, United States. Email: jcg219@math.rutgers.edu

[2] Department of Mathematics and Statistics, The College of New Jersey, Ewing, New Jersey 08628, United States

[3] Department of Electrical and Computer Engineering and Department of Bioengineering, Northeastern University, Boston, Massachusetts 02115, United States, and Laboratory of Systems Pharmacology, Harvard Medical School. Email: eduardo.sontag@gmail.com

References

[1] C. Holohan, S. Van Schaeybroeck, D. B. Longley, and P. G. Johnston, "Cancer drug resistance: an evolving paradigm," *Nature Reviews Cancer*, vol. 13, no. 10, pp. 714–726, 2013.

[2] A. O. Pisco, A. Brock, J. Zhou, A. Moor, M. Mojtahedi, D. Jackson, and S. Huang, "Non-darwinian dynamics in therapy-induced cancer drug resistance," *Nature communications*, vol. 4, 2013.

[3] J. Nyce, S. Leonard, D. Canupp, S. Schulz, and S. Wong, "Epigenetic mechanisms of drug resistance: drug-induced DNA hypermethylation and drug resistance," *Proceedings of the National Academy of Sciences*, vol. 90, pp. 2960–2964, 1993.

[4] J. Nyce, "Drug-induced DNA hypermethylation: a potential mediator of acquired drug resistance during cancer chemotherapy," *Mutation Research*, vol. 386, pp. 153–161, 1997.

[5] S. Sharma, D. Lee, B. Li, and M. e. a. Quinlan, "A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations," *Cell*, vol. 141, pp. 69–80, 2010.

[6] J. M. Greene, J. L. Gevertz, and E. D. Sontag, "Mathematical approach to differentiate spontaneous and induced evolution to drug resistance during cancer treatment," *J Clin Cancer Inform*, doi: 10.1200/CCI.18.00087, 2019.

[7] J. M. Greene, C. Sanchez-Tapia, and E. D. Sontag, "Control structures of drug resistance in cancer chemotherapy," *In Proc. 2018 IEEE Conf. Decision and Control*, pp. 5195–5200, 2018.

[8] ——, "Mathematical details on a cancer resistance model," *arXiv*, Sep 2018.

Reaction Rules for Whole-Cell Models

John A.P. Sekar[1], Arthur P. Goldberg[2], James R. Faeder[3] and Jonathan R. Karr[4]

***Abstract*— Predictive whole-cell models could enable simulation of complex cell phenotypes, rational cell design and precision medicine. To build whole-cell models, one must account for biochemical complexity in model representation and simulation. By compactly representing binding and covalent modification, the rule-based modeling approach has already enabled succinct pathway-scale models of signal transduction. Here, we present WC-Rules, a platform that builds on the rule-based approach and enables representing current models even more concisely. It also captures a broader range of molecular interactions such as steric effects, transcription and translation, and introduces new improvements to simulation.**

***Index Terms*— whole-cell models, rule-based models, network-free simulation, reaction rules**

I. Introduction

Whole-cell (WC) computational models that predict phenotype from genotype could potentially enable researchers to simulate complex phenotypes such as cancer, rationally design and engineer cells, and physicians to deliver precise, personalized care [1]. Unlike small models, where structural features that are not relevant to model scope can be abstracted away, a comprehensive model demands that we integrate and represent all structural and kinetic properties of biomolecules explicitly [1]. The reaction network approach is ill-suited for this purpose due to biochemical complexity (for example, combinatorial complexity of complexes [2], proteoform complexity [3], etc.). One promising approach is rule-based modeling [4], which describes binding and covalent modification reaction mechanisms using succinct *reaction rules*. This has already enabled building and simulating many pathway-scale signal transduction models [5]. Currently, the rule-based approach does not provide a sufficiently abstract representation that can be extended to accommodate other biochemical processes such as steric interactions, transcription and translation. This motivated us to develop WC-Rules, which generalizes and improves upon the current rule-based approaches, and can pave the way for WC model representation and simulation.

II. Methods

A. Model Representation

Current rule-based models follow a strict graph syntax [4]: complexes are composed from objects (molecules, sites) that have specific relations between them (molecules contain sites, sites bind other sites). Both molecules and sites are drawn from a pool of types defined by the modeler, and sites may be assigned a single label representing an internal state. We expand on this approach by introducing an extensible *bioschema*. Molecules, sites, bonds and other objects derive from a common `Entity` class that can be subclassed indefinitely. They can also be assigned any number of state attributes of various types as well as relations to each other. The subclassing allows more abstract representation of repetitive architecture (e.g., phospho-site adapter binding across multiple receptor types) as well as the novel ability to represent steric effects and sequence processivity using rules.

B. Model Simulation

When simulating a rule-based model, reaction events are sampled using a generalization of Gillespie's SSA algorithm applied to molecular graphs and graph transformation events [6]. A key component of the simulation algorithm is a pattern-matching engine that tracks how *patterns* in reaction rules map to the simulation state (e.g., a pattern such as 'unphosphorylated site' is used by a phosphorylation rule to track where phosphorylation can occur). Here, we implement pattern-matching using the *Rete-net* algorithm [7], which improves performance with significant reuse of partial matches. It also permits nested pattern definitions, which is a novel feature compared to current implementations. We also improve simulation efficiency by using the *Euler tour* technique [8] to track which molecules are in which complexes and a search path pruning technique to reduce matching overhead for self-symmetric patterns [9].

III. Conclusion

We expect these improvements will lead to even more concise forms of existing rules, which we demonstrate by applying WC-Rules to existing rule-based models. We also demonstrate new features such as rule-based models of transcription and translation. We believe these improvements will enable the integrated whole-cell models of the future.

This work is funded by NSF INSPIRE grant 1649014.

[1,2,4]Department of Genetics and Genomic Sciences, Mt. Sinai School of Medicine, New York. Email: [1]john.sekar@mssm.edu, [2]arthur.p.goldberg@mssm.edu, [4]jonathan.karr@mssm.edu

[3]Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh. Email: faeder@pitt.edu

References

[1] Goldberg A.P., *et al.* (2018) *Curr Opin Biotech* 51:97-102.
[2] Deeds E.J., *et al.* (2012) *PLOS ONE* 7(3): e32032.
[3] Harper J.W. and Bennett E.J. (2016) *Nature* 537:328-338.
[4] Chylek L.A., *et al.* (2014) *WIREs Syst Biol Med*, 6:13-36.
[5] Sekar J.A.P., *et al.* (2017) *PLOS Comp Biol* 13(11): e1005857.
[6] Suderman R., *et al.* (2018) *Bull Math Biol*.
[7] Forgy C. (1982) *Artificial Intelligence* 19:17-37.
[8] Henzinger M.R. and King V. (1995) *Proc. STOC '95*, p. 519.
[9] Katebi H., *et al.* (2010) *Proc. SAT '10*, p. 113-127.

Regulation of the ratio of surface area and volume synthesis in fluctuating environments

Handuo Shi[1], Yan Hu[1] and Kerwyn Casey Huang[1,2,3]

***Abstract*— Bacterial cells face constant environmental changes in their natural habitats, and therefore actively respond to those fluctuations. In the rod-shape bacterium *Escherichia coli*, it remains unclear how cells respond to continuous environmental changes and regulate cell shape homeostasis. In this work, we develop a time-delay model that fully captures the dynamics of surface area and volume synthesis in a laboratory batch culture. This model also predicts that surface area to volume ratio is robust to perturbations, but systematically changes with cell wall synthesis rates and translation rates, all of which have been experimentally verified. This work provides a quantitative description of cellular resource allocation.**

***Index Terms*— cell shape, surface area to volume ratio**

I. Introduction

Bacterial cells face constantly changing environments in their natural habitats, and cells adapt to their changing environments and stresses by altering their physiology. Cell shape is intrinsically linked to physiology, and theirfore cells actively alter both length and width at different growth conditions [1], [2].

Even in laboratory settings, bacterial cells grown in test tubes still face environmental changes: followed by the initial inoculation, cells resume growth when seeing fresh nutrients. After several hours of fast growth, the nutrients in the media become limiting and growth slows down. During this process, cell morphology actively changes, with larger width and length in exponential phase compared to stationary phase [3], [4].

Cell morphology in steady-state conditions are largely determined by the available nutrients [1], as well as how the cells allocate the available resources to expand their volumes and surface areas [2]. However, it remains unclear how this resource allocation paradigm re-adjusts when cells face constant changing environments.

In this work, we have modified the previous steady-state model on how cells balance surface area and volume growth [2] and include a time-delay component that accurately captures the dynamics of surface area to volume ratio (SA/V). The model also makes predictions on how SA/V changes upon different perturbations that are later experimentally verified.

[1]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA

[2]Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA

[3]Chan Zuckerberg Biohub, San Francisco, CA 94158

II. Results

A. Experimental observations

For a culture grown in a test tube, we took aliquots every 15 min to quantify cell morphology over time. The surface area to volume ratio (SA/V) dropped when cells resume growth, mainly due to an increase in cell width. When nutrient was depleted, cell growth slowed down and SA/V gradually increased to the initial value again. Such trends are universal across many rod-shaped organisms including Gram-positive and Gram-negative bacteria, fission yeasts, and are conserved at different culturing temperatures.

B. Model development

We found that the steady-state model [2] does not quantitatively capture the dynamics in SA/V changes. Considering that surface area and volume growth could change at different time scales, we incorporated a time-delay component in the original steady-state model, which yields a modified model that accurately predicts the dynamics of SA/V observed in experiments.

C. Model prediction

This model predicts that 1) the dynamics of SA/V is robust to genetic and chemical perturbations as long as cell growth and cell wall synthesis are not directly affected, 2) inhibiting cell wall synthesis reduces SA/V, and 3) increasing the time-delay constant causes larger ranges of SA/V, with lower SA/V in exponential phase and higher SA/V in stationary phase, compared to the control. All three predictions have been quantitatively verified experimentally, suggesting that our model accurately captures the key components of how cells allocate their limited resources to volumetric and surface area growth.

References

[1] F. Si, D. Li, S. E. Cox, J. T. Sauls, O. Azizi, C. Sou, A. B. Schwartz, M. J. Erickstad, Y. Jun, and X. Li, "Invariance of initiation mass and predictability of cell size in escherichia coli," *Current Biology*, vol. 27, no. 9, pp. 1278–1287, 2017.

[2] L. K. Harris and J. A. Theriot, "Relative rates of surface and volume synthesis set bacterial cell size." *Cell*, vol. 165, no. 6, pp. 1479–1492, 2016.

[3] H. Shi, A. Colavin, M. Bigos, C. Tropini, R. D. Monds, and K. C. Huang, "Deep phenotypic mapping of bacterial cytoskeletal mutants reveals physiological robustness to cell size," *Current Biology*, vol. 27, pp. 1–11, 2017.

[4] A. Colavin, H. Shi, and K. C. Huang, "Rodz modulates geometric localization of the bacterial actin mreb to regulate cell shape," *Nature communications*, vol. 9, no. 1, p. 1280, 2018.

Interacting Hopfield networks as a simple model for multicellular gene regulation

Matthew Smart[1] and Anton Zilman[2]

***Short Abstract* — The rapid development of single-cell RNA sequencing has led to widespread interest in dynamical modelling of cell state. A natural question that arises is how to encode identified cell types as attractors. Hopfield networks offer an elegant solution and have been used to describe reprogramming in individual cells. To capture cell-cell interactions, we propose to model interacting cells using a lattice of interacting Hopfield networks. We consider cell-cell interactions mediated by both ligand-receptor signalling and exosomes. We investigate under what conditions the single-cell attractors remain stable, and whether cell-cell interactions can facilitate the emergence of new stable single-cell states.**

I. Purpose

THE development of dynamical models capable of describing cellular state has become a topic of intense interest, driven by recent advances in experimental techniques such as high-throughput single-cell RNA sequencing (scRNA-seq) [1]. Clustered scRNA-seq data reveals stable transcriptomes associated with known cell types. A fundamental question is how to specify a dynamical system, in this case a gene regulatory network (GRN), for which the identified cell types correspond to attractors in gene expression space.

Hopfield networks [2] offer a simple solution to this problem. Given a matrix of cell type transcriptomes, they prescribe a matrix of gene-gene interactions which form the basis of an Ising model. This framework has been shown to recapitulate aspects of *in vitro* cellular reprogramming when gene expression noise and applied fields (which bias certain genes to be on or off) are tuned [3].

While this provides an exciting framework to dynamically model cell fate, its applications have been restricted to the level of single cells. A natural and theoretically interesting question is the following: given a dynamical model of intracellular gene regulation for which known cell types are fixed points, what happens when copies of the model interact via intercellular signalling? As this question relates to the situation multicellular organisms are faced with, answers will inform our understanding of development, homeostasis, and metaplastic diseases such as cancer.

II. Model

The basis for our framework is a lattice of communicating single-cell models. Each cell acts as a Hopfield network (Ising model) with gene regulation rules induced by encoding a set of identified cell types as stable expression states. The lattice state is then defined by the position and gene expression of each cell, and it evolves under stochastic gradient descent of the underlying multi-cell Ising Hamiltonian (using e.g. Glauber dynamics).

We consider two mechanisms to enable cell-cell communication via applied fields on the lattice: paracrine signalling, which allows a cell expressing signalling molecules (e.g. cytokines) to augment expression of specific genes in nearby recipient cells; and exosomes, universally secreted vesicles which can directly shuttle RNA and transcription factors between cells.

III. Results

We investigate under what conditions the encoded single-cell types remain stable, and whether cell-cell interactions can facilitate the emergence of new stable single-cell states within tissue. In our simulations, we consider arbitrary sets of low dimensional cell type vectors (~10 genes) to induce the single-cell Hopfield network, and we investigate the multicell behaviour under randomly sampled paracrine signalling rules. We focus on the deterministic (noiseless) limit.

As the cell-cell interaction strength is tuned, the multicell behaviour exhibits distinct phases. Importantly, we find that intermediate levels of either paracrine or exosome signalling can destabilize the encoded single-cell types. Whereas exosome signalling tends to promote spatial correlations in gene expression, leading to cell type "blurring", paracrine signalling can lead to complex tissue patterns composed of cell states that are unstable in the single-cell model.

Starting from initial conditions where the lattice consists of copies of a given encoded single-cell type at every point, it is possible, for example, for the cell-cell interactions to guide the system to maze-like structures made up of several "new" cell types which are not considered stable in the single-cell sense. Further, we find that geometric aspects of the patterns which form (e.g. width of stripes or maze paths) can be controlled by the interaction range.

Our framework allows for comprehensive analysis of multicellular transcriptomic state, and it provides new perspectives on cell type stability in the context of natural, pathological, and engineered cellular systems.

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC).

[1,2] Department of Physics, University of Toronto, Toronto, Canada.
E-mails: [1] msmart@physics.utoronto.ca, [2] zilmana@physics.utoronto.ca

References

[1] Rozenblatt-Rosen O, et al. (2017) The Human Cell Atlas: from vision to reality. *Nature* **550**, 451–453

[2] Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *PNAS* **79**(8), 2554-8

[3] Lang AH, et al. (2014) Epigenetic Landscapes Explain Partially Reprogrammed Cells and Identify Key Reprogramming Genes. *PLOS Comp Bio* **10**(8), e1003734

Continuous adaptation by the mTORC1-TFEB axis

Breanne Sparta[1], Nont Kosaisawe[1], Michael Pargett[1], and John Albeck[1]

The transcription factor EB (TFEB) is a central regulator autophagy genes. As with many transcription factors, TFEB is regulated through control of its subcellular localization. Multiple kinases, including GSK3β and mTORC1, phosphorylate TFEB to shift its equilibrium from the nucleus to the cytoplasm, enabling the cell to regulate autophagy in response to various energetic requirements. Because TFEB's localization is a function of multiple kinase activities and can be easily visualized, TFEB provides a model to understand how multiple upstream signals are integrated quantitatively within an individual cell. Using a live-cell reporter system for TFEB localization, we have tracked its response to a range of mitogenic and metabolic signals, including growth factors, amino acids, and glucose. We find that in individual cells, TFEB responds to sequential nutrient addition through incremental changes in localization. Using a series of reporters for TFEB and mTORC1, and the upstream regulators AKT and AMPK, we identify coherent, temporal fluctuations in the growth regulatory network in response to metabolic stress. Taken together, our observations support a model where continuous adaptation and temporal gating define the operating logic of the mTORC1-TFEB axis.

Acknowledgements: This work was funded in part by NIH NRSA grant 1-F31-GM120937-01A1.

[1] Department of Molecular and Cellular Biology, University of California Davis
E-mail: BSparta@ucdavis.edu, JGAlbeck@ucdavis.edu

Evolutionary

Natanael Sp

***Abstract*— The humoral response to an inf process of affinity maturation. In order to with good binding properties, B cells gather where they proliferate and acquire mutatio are selected to further multiply. This evolu fueled by somatic hypermutations (SHM) wh high frequency. We propose a repertoire-wid the properties of mutations on evolutionary help to understand the mutational landscap cells and assert the relevance of a phyloge identifying SHM hotspots.**

I. Background

Upon activation, B cells gather in the where they undergo rapid proliferat tate. The mutants are selected for the highe to a particular antigen. Mature B cells produce antibodies with desired properti eradicate the pathogenic threat. In many c evolves in parallel to the adaptive repertoi to an exciting co-evolutionary competitior

It is crucial that the cohort of high-affinity antibodies arrives on time and so the evolutionary process of affinity maturation needs to converge in a period of a few weeks. To achieve that, the evolution of B cells is fueled by an extremely high frequency of SHM, which changes the timescales by orders of magnitude as compared to the evolution of germline cells.

The rate of hypermutation is extremely heterogeneous across the sequence coding for the immunoglobulin molecules. It has been shown that to a large extent this can be explained by the context dependence [1]. Indeed, a mutability of a given position is determined by its nearest vicinity: some strings of nucleotides are much more likely to mutate than other. Effectively, a hierarchy of motifs emerges to guide the stochastic process of mutations.

It is also known that the context-dependence does not fully characterize the heterogeneity of the mutation profile. It has been observed that the mutations co-localize: there is a 4-fold enrichment in mutations at neighboring bases [2].

II. Results and discussion

The most important confounding factor in the study of hypermutations is one of selection. To get rid of this bias we perform an analysis of non-productive sequences i.e. the ones which the VDJ recombination rendered out-of-frame. The second confounding factor is known as phylogenetic bias.

This work was supported by grant ERCCOG no. 724208.

All at Laboratoire de Physique de l'Ecole Normale Supérieure, Paris.
[1] Email: spisak@ens.fr
[2] Equal contribution.

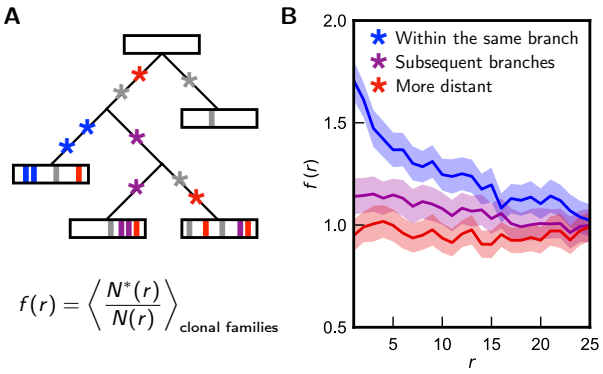


Fig. 1. A. Schema of hypermutation tree. B. Correlation function for the pairs of subsequent hypermutations in out-of-frame B cell lineages, data from [4]. $N(r)$ stands for the number of pairs distant by r positions.

For large enough families we reconstruct phylogenies using a maximum parsimony method. This approach allows us to establish the order at which the mutations happened as well as ensure that each mutation is counted once. It also makes it possible to properly identify the mutations in the highly diverse CDR region of the sequence. We introduce a correlation function which measures the enrichment in subsequent mutations i.e. the ones encoded in the same or neighboring branches. Traversing the phylogenetic trees of B cells we find that it is the consecutive mutations which tend to co-localize (Fig. 1B).

We have shown that an effective model of SHM needs to account not only for the context dependence but also for the relative order of mutations and their dependence on position. The mutation landscape of immunoglobulin-coding sequences is shaped by their motif composition as well as the co-localization of subsequent mutations. Disentangling the two effects remains a challenge.

References

[1] G. Yaari et al., "Models of somatic hypermutation targeting and substitution based on synonymous mutations from high-throughput immunoglobulin sequencing data," *Frontiers in Immunology*, vol. 4, 2013.
[2] Q. Marcou et al., "High-throughput immune repertoire analysis with IGoR," *Nature Communications*, vol. 9, 2018.
[3] D. Ralph and F. Matsen, "Likelihood-Based Inference of B Cell Clonal Families," *PLOS Computational Biology*, vol. 12, 2016.
[4] W. DeWitt et al., "A Public Database of Memory and Naive B-Cell Receptor Sequences," *PLOS One*, vol. 11, 2016.

Production of Protein-Complex Components Is Stoichiometric and Lacks General Feedback Regulation in Eukaryotes

James C. Taggart and Gene-Wei Li

Constituents of multi-protein complexes are required at well-defined relative levels. However, it remains unknown whether eukaryotic cells produce precise amounts of subunits, or rely on degradation to mitigate imprecise production. Here we quantified protein synthesis in diverse eukaryotes using ribosome profiling. Obligate components of multi-protein complexes are produced in proportion to their stoichiometry, indicating that their abundances are precisely set through synthesis. By systematically perturbing gene dosage in budding yeast, we found that negative feedback regulation buffering subunit synthesis is uncommon. These results reveal a principle of proteome homeostasis and highlight quantitative controls at every step in the central dogma.

Keywords **— Dosage compensation, feedback regulation, proportional synthesis, proteome homeostasis, ribosome profiling**

How precisely cells tune gene expression is a fundamental question in quantitative biology. A sufficient amount of each protein must be produced to satisfy the demand for its activities in the cell, and excess production can be wasteful. But to synthesize proteins at levels that are 'just enough' requires precisely tuned rates of transcription, translation, and mRNA decay, as well as selective pressure against overproduction. Elucidating the general principles underlying the set-points of protein synthesis rates is an important step towards understanding the design constraints of cells and the physiological impacts of gene expression perturbations.

Obligate subunits of multi-protein complexes are an abundant class of proteins whose minimally required levels are precisely defined. At steady state, the synthesis rate of each subunit must equal or exceed the minimum of its binding partners, but the benefit quickly diminishes when the subunit is made in excess. In bacteria, a systematic study of the rates of protein synthesis showed that nearly all obligate subunits of protein complexes are produced at rates proportional to their stoichiometry [1]. This principle, termed proportional synthesis, indicates that bacteria have evolved the ability to synthesize precise amounts of proteins, and that degradation of unassembled subunits is not the primary process controlling protein abundance.

In contrast to bacteria, it remains unclear whether precise proportional synthesis is a general strategy of gene expression in eukaryotes. Eukaryotic cells have much more elaborate protein degradation mechanisms, which could efficiently buffer against imprecise synthesis [2]. It has also been suggested that eukaryotes have lower selective pressure against wasteful production than bacteria, especially in multicellular organisms [3-4]. Many bacterial protein complexes are regulated by negative feedback that maintains expression levels [5-7], however negative feedback regulation is uncommon in yeast, although it remains to be determined whether subunits of complexes are exceptions to this rule [8-9]. A systematic analysis of obligate protein complexes and their synthesis rates is currently lacking.

In this work, we quantified the synthesis rates of eukaryotic protein complex subunits using ribosome profiling [10]. By manually curating a comprehensive list of stable and obligate protein complexes with well-defined stoichiometry in budding yeast, we found that the majority of complex subunits are synthesized at rates proportional to their stoichiometry. Proportional synthesis was also observed for large protein complexes in higher eukaryotic cells derived from primary tissues or whole animals. To investigate whether this precise synthesis is ensured by negative feedback loops, we systematically examined the response to altered gene dosage in yeast strains harboring an extra copy of single chromosomes. A twofold increase in synthesis rate was observed for the majority of complex subunits encoded by duplicated genes, including most ribosomal proteins. Together, these results show that eukaryotes evolved precise rates of protein synthesis, which are hard-coded in the genome without widespread feedback regulation.

Acknowledgements: This work was funded by NIH grant GM124732, GM007287. GL is a Searle Scholar, Pew Scholar, and Sloan Fellow. JT is an NSF Graduate Research Fellow

Department of Biology, Massachusetts Institute of Technology, 02142

REFERENCES

[1] Li GW, Burkhardt D, Gross C, Weissman JS (2014) Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell.* **157**, 624-35.

[2] Juszkiewicz S and Hegde RS (2018) Quality control of orphaned proteins. *Mol. Cell*. **71**, 443-57.

[3] Kafri M, Metzl-Ras E, Jona G, Barkai N (2016) The cost of protein production. *Cell Rep.* **14**, 22-31.

[4] Lynch M and Marinov GK (2015) The bioenergetics costs of a gene. *PNAS.* **112**, 15690-5.

[5] Freedman LP, Zengel JM, Archer RH, Lindahl L (1987) Autogenous control of the S10 ribosomal protein operon of *E. coli*: genetic dissection of transcriptional and posttranscriptional regulation. *PNAS*. **84**, 6516-20.

[6] Mattheakis LC and Nomura M (1988) Feedback regulation of the spc operon in *E. coli*: translational coupling and Mrna processing. *J. Bact.* **170**, 4484-92.

[7] Shen-Orr SS, Milo R, Mangan S, Alon U (2002) Network motifs in the transcriptional regulation network of *E. coli*. *Nat. Genet.* **31**, 64-8.

[8] Ishikawa K et al. (2017) Post-translational dosage compensation buffers genetic perturbations to stoichiometry of protein complexes. *PLoS Genet.* **13**, e1006554.

[9] Springer M, Weissman JS, Kirschner MW (2010) A general lack of compensation for gene dosage in yeast. *Mol. Syst. Biol.* **6,** 368.

[10] Taggart JC and Li GW (2018) Production of protein-complex components is stoichiometric and lacks general feedback regulation in eukaryotes. *Cell Systems*. **7,** 580-9.

An Off-Lattice Spatial Next-Reaction Algorithm

Marcus Thomas[1], Russell Schwartz[2]

***Short Abstract* — We develop a novel generalization of the Gillespie algorithm to model bimolecular association reactions under diffusive motion. A rejection sampling method is presented for exact and efficient sampling of reaction locations in spatially heterogeneous regions. We explore our method's ability to capture an important subset of cellular self-assembly systems for which classic SSA and spatial Gillespie are too simplified but standard Green's Function Reaction Dynamics (GFRD), lattice-based reaction-diffusion methods, and Brownian Dynamics (BD) are too inefficient to accurately model the phenomena in realistic biological parameter ranges.**

***Keywords* — Off-lattice Reaction Diffusion, Particle Based Discrete Event Simulation, Gillespie Algorithm**

I. Background

THE ability of collections of molecules to spontaneously assemble into large functional complexes is central to nearly all cellular processes. Yet the fine details of complex self-assembly processes normally cannot be directly experimentally observed due to their small scale and rapid dynamics. Intractability of experimental approaches is particularly acute for understanding self-assembly in vivo, which may operate quite differently from purified in vitro models due to such effects as spatial confinement, macromolecular crowding, and influences of extrinsic cellular factors. Simulation has thus proven a valuable adjunct to experimental methods, facilitating interpretation of experimental data and inferences about experimentally unobservable aspects of self-assembly [1].

Many simulation methodologies have been used to model self-assembly dynamics --- including approaches based on Differential Equations (DE), Gillespie's SSA [2,3], Brownian Dynamics (BD), Green's Function Reaction Dynamics (GFRD) [4-6], lattice reaction-diffusion [7], and variations thereon --- each offering tradeoffs with respect to the phenomena they can model, their computational tractability, and the difficulty of fitting them to experimental data.

II. Methods

We explore a generalization of the SSA which assumes that reactions occur in a spatially heterogeneous environment that must be explicitly modeled. Next reaction times are computed using time-dependent propensity functions calculated from the diffusion coefficients of both particles and their last known positions. Unlike GFRD, our method explicitly models densities of particle positions in overlapping diffusion spheres to allow efficient and exact sampling of reaction times. Given two particles and their next reaction time, reaction locations are determined from the intersection of their probability distributions, using a rejection sampling method to sample exactly from this complicated distribution.

III. Results and Conclusion

A version of the simulator with periodic boundary conditions has been implemented in Matlab. The runtime is strongly dependent on the calculation of potential bimolecular events, less so on potential unimolecular and position-only-update events, and on event execution. There exists a tradeoff between the number of potential events considered (e.g. only computing bimolecular wait times for "nearby" partners) and the duration of the corresponding wait times. Future work will attempt to minimize the number of unnecessary potential events considered through adaptive definitions of "nearby". We also plan to evaluate the algorithm's runtime efficiency on the benchmarks described in [7], and to infer diffusion coefficients from assembly structure.

References

[1] Thomas, M., & Schwartz, R. (2018). A method for efficient Bayesian optimization of self-assembly systems from scattering data. *BMC Systems Biology*, 12(1), 65.

[2] Gillespie, D. T. (1992). A rigorous derivation of the chemical master equation. *Physica A: Statistical Mechanics and its Applications*, *188*(1-3), 404-425.

[3] Anderson, D. F. (2007). A modified next reaction method for simulating chemical systems with time dependent propensities and delays. *The Journal of Chemical Physics*, *127*(21), 214107.

[4] Mathai, A. M., & Provost, S. B. (1992). *Quadratic forms in random variables: theory and applications*. Dekker.

[5] van Zon, J. S., & Ten Wolde, P. R. (2005). Simulating biochemical networks at the particle level and in time and space: Green's function reaction dynamics. *Physical Review Letters*, *94*(12), 128103.

[6] Sokolowski, Thomas R., et al. "eGFRD in all dimensions." *The Journal of chemical physics* 150.5 (2019): 054108

[7] Chew, Wei-Xiang, et al. "Reaction-diffusion kinetics on lattice at the microscopic scale." *Physical Review E* 98.3 (2018): 032418.

This work was funded by NIH grant #T32EB009403 and the Computational Biology Dept of Carnegie Mellon University.

[1]Computational Biology Department, Carnegie Mellon University.
[2]Department of Biological Sciences, Carnegie Mellon University.
Email: marcust@andrew.cmu.edu, russells@andrew.cmu.edu

Protein quality control tunes fitness landscapes

Samuel Thompson[1,2], Christine Ingle[3], Kimberly Reynolds[3], and Tanja Kortemme[1]

***Short Abstract* — The cellular context determines how protein sequences are optimized by evolution, and one large determinant is protein homeostasis. Here, we explored how the *in vivo* fitness landscape of single point mutations to E.coli dihydrofolate reductase (DHFR) is shaped by constraints for protein homeostasis imposed by Lon protease. Under strict selection for DHFR activity, we observed mutations throughout DHFR that improved fitness. These beneficial mutations are generally repressed by Lon protease activity. From these observations, we propose a model in which the fitness landscape is tuned by the relative levels of refolding and degradation activities in the cell.**

***Keywords* — evolution, activity-stability tradeoffs, deep mutational scanning, fitness landscape, protein homeostasis, enzyme function, Lon protease, DHFR**

I. Background

EVOLUTION optimizes the genetic sequence to the fitness requirements of the organism. Performing the reverse calculation to predict phenotype starting from genotype is a major scientific goal, but even predicting the functional impact of individual mutations remains a significant challenge. At the protein level, activity-stability trade-offs are an intrinsic property of protein sequences[1], but mutational impact in the cellular environment ultimately determines the fitness of a protein. A more comprehensive understanding of the cellular constraints on protein function is needed. In this work, we utilized deep mutational scanning[2, 3] to measure the *in vivo* fitness impacts of single point mutations to an essential enzyme in E. coli, DHFR[4]. We measured fitnesses in the presence and absence of Lon protease, a key regulator of protein quality control[5, 6].

II. Results

Under stringent selection for DHFR activity in the absence of Lon protease, ~15% of all single point mutations to DHFR produce a beneficial phenotype, and at least one beneficial mutation exists for nearly every residue in DHFR. *In vitro* kinetics measurements on selected beneficial mutants show near-WT or increased DHFR activity. Beneficial mutations are not consistent with predictions of mutational tolerance from models based on either homology (MSAs, PAM substitution matrices[7]) or biophysics (Rosetta[8, 9]). From structural analyses, we observe that deleterious mutations cluster near the site of catalysis, while the likelihood that a mutation will be beneficial increases with distance from the site of catalysis. Based on the identity of the beneficial mutations, we infer that they introduce packing defects. Consistent with this observation, the activity of Lon protease generally dampens or suppresses the beneficial mutations.

III. Conclusions

We show here that Lon protease tunes activity-stability trade-offs in deep mutational scanning of E. coli DHFR. From structural mapping of mutational impacts, we more broadly interpret that the DHFR fitness landscape in the cellular environment is shaped by multiple antagonistic pressures, including catalysis, product release, product inhibition, and stability. We consider these observations in the context of previous studies showing the mutational buffering capacity of molecular chaperones (e.g. Hsp90[10] and GroEL[11]), and from this we propose a model in which molecular chaperones and proteases are antagonistic forces on the fitness landscape. In this evolutionary model, access to destabilizing gain-of-function mutations is enabled by increased refolding activity and is restricted by increased degradation activity.

References

1. Shoichet, B.K., et al., *A relationship between protein stability and protein function.* Proc Natl Acad Sci U S A, 1995. **92**(2): p. 452-6.
2. Rubin, A.F., et al., *A statistical framework for analyzing deep mutational scanning data.* Genome Biol, 2017. **18**(1): p. 150.
3. Fowler, D.M., J.J. Stephany, and S. Fields, *Measuring the activity of protein variants on a large scale using deep mutational scanning.* Nat Protoc, 2014. **9**(9): p. 2267-84.
4. Reynolds, K.A., R.N. McLaughlin, and R. Ranganathan, *Hot spots for allosteric regulation on protein surfaces.* Cell, 2011. **147**(7): p. 1564-75.
5. Gur, E. and R.T. Sauer, *Recognition of misfolded proteins by Lon, a AAA(+) protease.* Genes Dev, 2008. **22**(16): p. 2267-77.
6. Bershtein, S., et al., *Protein quality control acts on folding intermediates to shape the effects of mutations on organismal fitness.* Mol Cell, 2013. **49**(1): p. 133-44.
7. Mount, D.W., *Comparison of the PAM and BLOSUM Amino Acid Substitution Matrices.* CSH Protoc, 2008. **2008**: p. pdb.ip59.
8. Smith, C.A. and T. Kortemme, *Predicting the tolerated sequences for proteins and protein interfaces using RosettaBackrub flexible backbone design.* PLoS One, 2011. **6**(7): p. e20451.
9. Alford, R.F., et al., *The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design.* J Chem Theory Comput, 2017. **13**(6): p. 3031-3048.
10. Queitsch, C., T.A. Sangster, and S. Lindquist, *Hsp90 as a capacitor of phenotypic variation.* Nature, 2002. **417**(6889): p. 618-24.
11. Tokuriki, N. and D.S. Tawfik, *Chaperonin overexpression promotes genetic variation and enzyme evolution.* Nature, 2009. **459**(7247): p. 668-73.

Acknowledgements: This work is supported by NSF MCB-1615990 and NSF XSEDE allocation MCB-160192. Samuel Thompson is supported by the National Science Foundation Graduate Research Fellowship Program. Equipment for the *in vivo* selection was provided by Rama Ranganathan. The authors are grateful to Carol Gross for her insightful suggestions.

[1]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco. E-mail: samuel.thompson@ucsf.edu
[2]Graduate Program in Biophysics, University of California, San Francisco
[3]Green Center for Systems Biology and Department of Biophysics, University of Texas Southwestern

Topology-Dependent Interference of Synthetic Gene Circuits by Growth Feedback

Rong Zhang[1,#], Jiao Li[1,2,#], Juan Melendez-Alvarez[1], Xingwen Chen[1], Patrick Sochor[1], Qi Zhang[1], Hanah Goetz[1], Tian Ding[2], Xiao Wang[1,*], and **Xiao-Jun Tian**[1,*]

***Short Abstract* — Growth-mediated feedback between synthetic gene circuits and host organisms leads to various emerged behaviors, including innate growth bistability and increased ultrasensitivity. However, the adverse impact of growth feedback on gene circuits remains unexplored. Here, we found that the effects of growth feedback on the functional perturbations of gene circuits depend on the network topology. Specifically, the memory of a self-activation circuit is lost due to the fast growth of host cells. Decoupling of growth feedback reveals its hysteresis property in a broad range. Interestingly, the toggle switch circuit is more refractory to the growth feedback. The underlying principle is demonstrated by modeling the interplay between microbial growth and the gene circuit. Our results reveal a topology-dependent mechanism underlying the functional perturbation of gene circuits by growth-mediated feedback.**

***Keywords* — Bistability, Memory loss, Growth feedback.**

I. INTRODUCTION

The synthetic gene circuits are inevitably coupled with the growth feedback, in which the gene circuit affect the cell growth and the cell growth affect the expression of the genes in the circuits [1, 2]. Growth-mediated feedback can endow the synthetic gene circuits with various emerged properties. For example, a non-cooperative positive autoregulatory system, when coupled with growth-mediated feedback, gained significantly increased effective cooperativity and thus resulted in bistability [3, 4]. Toxin cooperativity can also be induced in the multiple toxin-antitoxin system with several growth-mediated feedback [5]. However, the designed gene circuits are disguised by the growth-mediated feedback and its true function may not be evaluated clearly. For example, if one gene circuit can serve as a bistable switch when coupled with growth-mediated feedback but it may not work in the stationary phase, or vice versa. Given that wild bacterial in nature stay in the stationary phase most of their lifetime [6], it is very interesting to test our synthetic circuits in this overlooked physiological state and evaluate their functions when the growth-mediated feedback is decoupled.

II. RESULTS

In this work, we tested whether the circuit-host interaction negatively affect circuits and disguised their true behaviors. We built several memory gene circuits and tested their memory maintenance in different cell growth conditions guided by mathematical modeling**.** First, we built one mathematical model and fitted the parameter with quantitative experiments data of the gene circuit. Based on the theoretic analysis, we predicted that this circuit is bistable. However, further experiments with the dilution of the GFP ‘ON' and ‘OFF' cells into fresh medium with different concentrations of stimulus showed no hysteresis, indicating inconsistence between in vitro and in silico experiments and the memory loss of circuit in vitro.

To find out the underlying reason for the memory loss, we systematically measured the dynamics of the cell after dilution. We found that GFP ‘ON' cells completely switched off 3 hours after dilution and then reached to different steady levels of GFP based on the concentration of stimulus. Thus, dilution into fresh medium inhibits the expression of gene circuit and growth-mediated feedback disguised the bistability of the gene circuit.

To decouple the growth-mediated feedback loop, we developed several dilution protocols and found that GFP level of the ‘ON' cells were maintained even stimulus was completely removed, indicating that the system functions as an irreversible bistable switch in absence of cell growth and the prediction of the model is verified. Thus, decoupling of the growth-mediated feedback reveals the bistability from the gene circuit.

In order to study whether the growth feedback also affects other types of gene circuits, we further tested our protocols for the toggle switch and found that the impact of growth feedback on toggle switch is minimal. After integrating the cell growth into our model, the underlying mechanism for this difference was demonstrated and was further verified with microfluidic device.

III. CONCLUSION

Here we found that circuit-host interaction also has an adverse impact on gene circuits by disguised their true behaviors, which can be revealed by decoupling this interaction.

REFERENCES

[1] Klumpp, S. and T. Hwa. Curr Opin Biotechnol, 2014. 28: 96-102.
[2] Klumpp, S., Z. Zhang, and T. Hwa, Cell, 2009. 139: 1366-75.
[3] Tan, C., P. Marguet, and L. You, Nat Chem Biol, 2009. 5: 842-8.
[4] Nevozhay, D., et al. PLoS Comput Biol, 2012. 8(4): p. e1002480.
[5] Feng, J., et al. Proc Natl Acad Sci U S A, 2014. 111: 544-9.
[6] Gefen, O., et al. Proc Natl Acad Sci U S A, 2014. 111: 556-61.

Acknowledgements: This work was funded by Arizona State University Start-Up Grant, NSF grant (EF-1921412) (to X-J.T.)

[1]School of Biological and Health Systems Engineering, Arizona State University, Tempe, AZ 85281, United States of America. E-mail: xiaojun.tian@asu.edu
[2]Department of Food Science and Nutrition, Zhejiang University, Hangzhou, Zhejiang, China

Exploring the effects of signaling gradients on the dynamics of the segmentation clock oscillations

Takehito Tomita[1], Volker Lauschke[1,2], and Alexander Aulehla[1]

***Short Abstract* — Vertebrate segmentation occurs as a periodic process underlay by oscillatory genes expressed coherently across the tissue. Notably, target genes of the Notch pathway show expression as traveling phase waves in the presomitic mesoderm (PSM). In these genes, we also observe a period gradient across the tissue, with the fastest oscillation at the posterior. This period gradient can account for the buildup of phase shift and therefore traveling waves. However, what controls this period gradient is totally unknown. Here we show that the modulation of signaling molecules affect the period gradient, and propose a theoretical framework to explain such results.**

***Keywords* — Segmentation clock, signaling gradients, oscillations, somitogenesis**

I. Period gradient in the PSM

In the mouse model, a somite is made every two hours. Corresponding oscillatory gene activity are repeated for every segmentation event; i.e., when observing certain target genes of the Notch pathway, a traveling wave of expression is seen every two hours. However, quantification of such gene expression by live reporters reveal that cells show oscillations with different periods, depending on their position along the AP axis. The period is approximately 130 minutes in the posterior tip and 170 minutes in the anterior, comprising a period gradient across the PSM [1]. Interestingly, when the posterior tip of the developing tail is cut and plated face down on a glass surface, this period gradient reemerges in the central periphery axis, and again creates traveling waves [2]. Furthermore, when cells from the PSM are randomized and plated as an aggregate, multiple small foci form de novo, and self-organize this period gradient [1]. While the period gradient is observed in these multiple contexts, how it is controlled or more generally how the dynamics of these oscillations are controlled is not fully understood, and is a key question in the field.

II. Experimental approach and results

In our approach, we quantify the oscillatory activity of the segmentation clock genes using highly dynamic fluorescent reporter lines [3]. By culturing and live imaging tail bud explants in vitro on a 2D surface [2], we can reliably quantify the emerging period gradient. The culture conditions can be easily modified to contain various amounts of signaling molecules or its inhibitors.

We ask the question whether signaling gradients that exist within the PSM play a role in modulating oscillation period, thereby affecting the period gradient. FGF8 is one such molecule that is found as a gradient in the PSM, high in the posterior and low in the anterior [4]. Indeed, our in vitro assays reveal that modulation of FGF8 levels affect oscillation dynamics, and the emergence of the period gradient. Specifically, the time necessary to establish a period gradient increases with higher FGF8, and the shape of the period gradient becomes shallower. Correspondingly, the wave dynamics seen are changed in a quantifiable manner.

III. Conclusions

We build a phenomenological model to explain these experimental results and provide a novel framework of how the period gradient may be controlled in the PSM. Our goal is to then test predictions derived from the model experimentally for further validation and refinement, and to ultimately understand the role of signaling gradients in controlling the dynamics of the segmentation clock oscillations.

References

[1] Tsiairis CD, Aulehla A (2016). Self-Organization of Embryonic Genetic Oscillators into Spatiotemporal Wave Patterns. Cell 164, 656-667

[2] Lauschke VM, et al. (2013). Scaling of embryonic patterning based on phase-gradient encoding. Nature 493, 101-105.

[3] Aulehla A, et al. (2008). A beta-catenin gradient links the clock and wavefront systems in mouse embryo segmentation. Nat. Cell Biol. 10, 186-193

[4] Dubrulle J, Pourquié O (2004). *fgf8* mRNA decay establishes a gradient that couples axial elongation to patterning in the vertebrate embryo. Nature 427, 419-422

Acknowledgements: This work was funded by European Research Council (ERC) Grant. This work has been supported by Laboratory Animal Resources (LAR), Advanced Light Microscopy Facility (ALMF) EMBL Heidelberg.

[1]Developmental Biology Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Germany E-mail: takehito.tomita@embl.de, aulehla@embl.de

[2]Present address: Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden E-mail: volker.lauschke@ki.se

A Software for Reaction–Diffusion Simulations of (Non–)Equilibrium Self–Assembly

Matthew J. Varga[1] and Margaret E. Johnson[1]

***Abstract*— Self–assembly is a critical step in many biological processes, including endocytosis and virion formation. Most existing computational methods to study self-assembly are limited by their lack of spatial resolution or in the ability to reach biologically–relevant timescales. We detail generalized software for performing reaction–diffusion simulations of self-assembly through recently developed algorithms that include rigid-body structure and interface–level resolution. This software accurately reproduces the kinetics, and equilibrium, of association in an efficient manner, which we demonstrate through the simulation of several model biological processes.**

***Index Terms*— reaction diffusion, self–assembly, simulation, endocytosis, modeling**

I. Purpose

DEVELOPING computational methods to model cell–scale processes is an active field of research. These tools provide spatial and temporal resolution unavailable to experiments, and can predict mechanistic detail which can subsequently be tested experimentally. Several methods have been developed to solve for or simulate the time–evolution of species in biological systems, ranging from concentration–based methods, such as reaction–diffusion master equation, to single–particle methods, such as that used in the Smoldyn package [1]. However, most of these existing methods lack either the spatial resolution, to account for factors such as nonhomogeneity in the cell, or temporal resolution necessary to simulate self–assembly. We detail here software using the free–propagator reweighting (FPR) reaction–diffusion algorithm, previously developed in our lab [2], [3], to simulate biological self–assembly. This method is unique in its ability to couple spatial self–assembly simulations on biologically–relevant timescales with non–equilibrium events, such as cooperativity.

II. Software

Each simulated particle (molecule, protein, etc.) is represented as a set of center of mass and interface coordinates and are propagated with positions sampled from the free–diffusion propagator. As reaction probabilities are agnostic to orientation in this algorithm [4], corrected for by an effective diffusion constant, a preset geometry is enforced upon association of two proteins to prevent the formation of arbitrary structures. This associated complex geometry is defined by several angles and vectors within and between the bound proteins.

Simulation parameters and reactions are provided by the user in a simplified BioNetGen language (BNGL) syntax [5], wherein reactions are represented as rules, along with angles for rigid-body association. This rule–based reaction syntax also allows for coupling to non–equilibrium phenomena.

III. Biological System Models

We demonstrate this software through several examples of equilibrium and non–equilibrium model systems.

A. Clathrin–mediated endocytosis

Ubiquitous in eukaryotic cells, clathrin–mediated endocytosis is a self–assembly process which is the primary method for cargo uptake from the extracellular environment. Several clathrin–mediated endocytosis systems were simulated, providing us with detail on

- the dependence of clathrin assembly on K_D,
- the effect of membrane localization on adaptor protein–facilitated clathrin assembly,
- the effects of irreversible lipid phosphorylation on clathrin assembly on the membrane.

B. Genetic oscillator

We reproduce oscillations in protein expression driven by a simple gene activator and repressor model [6].

IV. Conclusion

We developed software utilizing the free–propagator reweighting reaction–diffusion algorithm for simulating equilibrium and non–equilibrium self–assembly with spatiotemporal resolution. This software has been demonstrated through simulations of several biological model systems, and represents a unique and powerful tool for cell–scale simulations.

[1]Department of Biophysics, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218 Email: margaret.johnson@jhu.edu

References

[1] S. S. Andrews, N. Addy, R. Brent, and A. Arkin, *PLoS Computational Biology*, vol. 6, p. e1000705, 2010.
[2] M. E. Johnson and G. Hummer, *Physical Reviews X*, vol. 4, no. 3, p. 031037, 2014.
[3] O. N. Yogurtcu and M. E. Johnson, *Journal of Chemical Physics*, vol. 143, no. 8, p. 084117, 2015.
[4] M. E. Johnson, *The Journal of Physical Chemistry B*, vol. 122, no. 49, pp. 11 771–11 783, 2018.
[5] M. L. Blinov, J. R. Faeder, B. Goldstein, and W. S. Hlavacek, *Bioinformatics*, vol. 20, no. 17, pp. 3289 – 3291, 2004.
[6] J. M. G. Vilar, H. Y. Kueh, N. Barkai, and S. Leibler, *Proceedings of the National Academy of Sciences*, vol. 99, no. 9, pp. 5988–5992, 2002.

From useful metaphor to quantitative framework: a unified view of Waddington landscapes

John J. Vastola[1], William R. Holmes[1]

***Abstract*— Waddington's metaphor for development—cells as balls rolling around a rugged landscape—is difficult to make mathematically precise. In a probabilistic view of the problem, it is not generally possible to define a landscape that accurately captures both (i) relative state occupancy information and (ii) cell state transition information simultaneously. As a partial remedy, we discuss a comprehensive way to think about landscapes, which landscapes answer which biological questions, and the qualitative importance of incorporating noise. We offer a unifying general definition of landscapes, and discuss how existing proposals satisfy it. Finally, we illustrate our ideas with examples from development and reprogramming.**

***Index Terms*— landscape, stochastic models, gene regulation, development, reprogramming, path integral, Langevin equation, stochastic differential equation, chemical master equation**

I. Purpose

Though Waddington's epigenetic landscape has historically been a useful mental model for thinking about development and differentiation, it has been hard to make mathematically precise. Many different quantitative attempts to realize Waddington's idea exist in the literature, but these are inequivalent in general [1]. What is the 'right' notion of a landscape, and why has it been so hard to pin down?

Waddington imagined landscapes to answer two questions: (i) how much more likely is one state than another state?; and (ii) what path will a cell take through state space when it transitions between states? Biologically, these questions correspond to (for example) thinking about the efficiency of a reprogramming protocol, or the optimal phase space path for reprogramming one cell type to another.

As we will discuss, the central difficulty in providing a mathematically precise description of a landscape is that no single landscape can simultaneously answer both questions. Still, given the proliferation of proposed landscapes, there is a need to better understand their relationships and why some candidates seem more useful than others.

II. Results

We broadly classify landscapes as either global (conveying relative state occupancy information) or local (conveying transition path information), and present new and general definitions for both kinds of landscapes that emphasize their topological relationship with steady state probability distributions and transition probability distributions respectively. These definitions do not assume that the system is defined via a particular mathematical model (like Langevin equations or a chemical master equation), but do require that the system is fundamentally probabilistic.

Our definition particularly clarifies how to think about local landscapes, since it makes clear that they are in general associated with a base point and a timescale. In other words, the local landscape you construct will qualitatively depend on (i) which state the transitions you care about begin from, and (ii) the finite length of time of your experiment. This picture unifies the many different definitions of local landscapes in the literature, and makes the connection to Doi-Peliti/path integral literature [2] more transparent.

We discuss the relationship between our definition and many landscapes from the literature, and show that some candidates are only landscapes (in the sense of our definition) in certain limits. For example, the normal decomposition landscape [3] and the Freidlin-Wentzell landscape [1] are only landscapes in the small, symmetric additive noise limit. We discuss how, in some sense, the existence of many different kinds of landscapes reflects the need for more or less information about the underlying probability distributions.

Finally, we describe the consequences of these definitions for biologists intending to use landscapes as a practical tool, and illustrate our points using examples from development and reprogramming.

III. Conclusion

Taking a probabilistic perspective facilitates a view of landscapes that is model-independent, emphasizes how different kinds of intrinsic noise can qualitatively change landscape structure, and more intimately connects with what is experimentally observable than dynamical systems-based approaches. Through this lens, it becomes clear that there is not one Waddington landscape, but many; these landscapes can be broadly classified as either global or local, depending on whether they convey state occupancy information or transition path information.

This work was funded by NSF grant DMS-1562078.

[1]Department of Physics and Astronomy, Vanderbilt University. Email: john.j.vastola@vanderbilt.edu, william.holmes@vanderbilt.edu

References

[1] P. Zhou and T. Li, "Construction of the landscape for multi-stable systems: Potential landscape, quasi-potential, a-type integral and beyond," *J Chem Phys*, vol. 144, no. 9, p. 094109, 2016.

[2] M. F. Weber and E. Frey, "Master equations and the theory of stochastic path integrals," *Rep. Prog. Phys.*, vol. 80, no. 4, p. 046601, 2017.

[3] J. Zhou *et al.*, "Quasi-potential landscape in complex multi-stable systems," *J Royal Soc Interface*, vol. 9, no. 77, pp. 3539–3553, 2012.

Collective Sensing by Cell Populations with Feedback and Communication

Michael Vennettilli[1], Amir Erez[2] and Andrew Mugler[1]

***Abstract*— Cells sense their environment with remarkable precision, and this precision can be enhanced by cell-cell communication. However, most theoretical investigations of this effect have analyzed linear networks, while cells use nonlinear feedback to internally amplify sensed signals. We use a minimal stochastic model to investigate the interplay of sensing, communication, and feedback. We find that feedback can induce a critical transition with long-range order. We investigate the associated sensing tradeoff: on the one hand, we expect long-range order to enhance communication; on the other hand, fluctuations become large at criticality, so order may come at the cost of precision.**

***Index Terms*— cellular sensing, cell-cell communication, feedback, criticality**

Cells need to sense their environment. For example, bacteria can detect the presence of nutrients or harmful compounds in their environment and use this information to navigate towards more habitable regions. We also know that cells can communicate information to each other through exchanging molecules. This exchange can be used to communicate sensory information about the environment, which can improve sensory precision by performing a spatial average [1]. This communication is also necessary to coordinate collective behaviors between cells.

Much of the work on this topic has looked at linear models, but cells use positive and negative nonlinear feedback to modulate signals. One system with sensing, communication and feedback is the Bicoid-Hunchback system in the developing *Drosophila* embryo. A Bicoid concentration gradient is used to differentiate the front and back of the embryo. This is partially achieved by activating the Hunchback gene. The gene exhibits positive autoregulation, and Hunchback proteins are thought to be exchanged between cells [2]. This exchange is able to improve sensory precision by reducing the noise due to bursts [3]. Another system with these three features is the quorum sensing network in *A. fischeri* and *V. harveyi*. This network is used to detect the density of neighboring cells and coordinate a transition to a collective bioluminescent state. LuxI produces an autoinducer that is released into the environment. The autoinducer can bind to receptors on the surface of other cells, which activates LuxR, which in turn produces more LuxI. Experiments have shown that this circuit exhibits bistability and hysteresis [4]. This has been explained through LuxR exhibiting positive autoregulatory feedback.

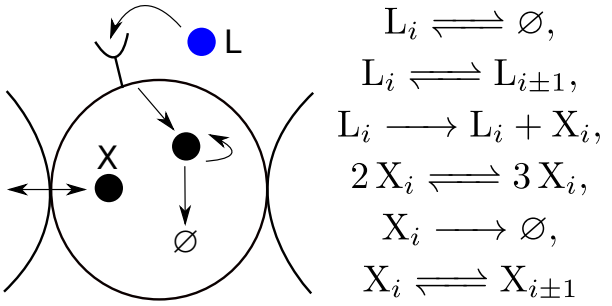


Fig. 1. Model schematic and reactions. A ligand L is sensed by cell i, causing production of an internal readout X that is subject to autoregulatory feedback and is communicated between cells. The third reaction describes sensing, the fourth describes non-linear feedback, and the last describes cell-cell communication.

We formulated a model which attempts to capture the basic features of these systems, namely sensing, communication, and feedback. Our model is a variant of Schlögl's second model [5]. We have a lattice where each site corresponds to a cell and its immediate environment. Each site can have an unbounded number of molecules. There are two molecular species: an external ligand L and an internal readout X. The precise reactions used are shown in Fig. 1.

The positive and negative feedback introduce a bifurcation in the system's behavior. The bifurcation point has the scaling properties of a second-order phase transition in the mean field Ising model's universality class. Using a combination of stochastic calculus, Gillespie simulations, a Ginzburg-Landau description, and information theory, we probe the implications of being at this critical point for sensing and its interplay with cell-cell communication.

References

[1] S. Fancher and A. Mugler, "Fundamental limits to collective concentration sensing in cell populations," *Phys. Rev. Lett.*, vol. 118, p. 078101, Feb 2017.

[2] J. Treisman and C. Desplan, "The products of the drosophila gap genes hunchback and krüppel bind to the hunchback promoters," *Nature*, vol. 341, no. 6240, p. 335337, 1989.

[3] T. Erdmann, M. Howard, and P. R. ten Wolde, "Role of spatial averaging in the precision of gene expression patterns," *Phys. Rev. Lett.*, vol. 103, p. 258101, Dec 2009.

[4] J. W. e. a. Williams, "Robust and sensitive control of a quorum-sensing circuit by two interlocked feedback loops," *Molecular Systems Biology*, vol. 4, no. 1, 2008.

[5] F. Schlögl, "Chemical reaction models for non-equilibrium phase transitions," *Zeitschrift für Physik*, vol. 253, no. 2, p. 147161, 1972.

This work was funded by Simons Foundation grant 376198.

[1]Department of Physics and Astronomy, Purdue University, West Lafayette, Indiana 47907, USA. Email: mvennett@purdue.edu, amugler@purdue.edu

[2]Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA.

RNA-Seq Reveals Regulatory Pathways in a New Kind of Persister Cell Phenotype

Huijing Wang[1,3], GW McElfresh[1], Nishantha Wijesuriya[2], Andrew D. Hecht[2], and Christian J. Ray[1,2,3]

***Short Abstract* — Cell regulatory networks with numerous entangled interactions create phenotypes that are complex, with many genes contributing to the overall cell fate. Advances in sequencing technology provide hope that we can identify principles of complex traits with high throughput data. In our previous work, we identified a novel antibiotic-tolerant persister phenotype in *E. coli* induced by metabolic excess, contradicting the belief that persisters arise primarily from starvation [1]. RNA sequencing (RNA-Seq) of this strain in various conditions allowed us to perform transcriptome profiling for gene, gene ontology, and pathway analysis. Comparing starving and metabolically toxified cells to moderate conditions, we find differential expression of ribosomal proteins, ATP synthesis genes, stress responses, and cell envelope genes. Our results suggest a multifaceted response centered on feedback between the cell cycle and cell wall synthesis.**

***Keywords* — persister, RNA sequencing, differential expression analysis, gene regulation.**

I. Introduction

BACTERIA are considered simple single celled organisms, but they contain high dimensional interaction networks that biologists still struggle to understand. Among the more pressing questions with medical relevance is how they survive stressful conditions, such as the famous antibiotic-tolerant persister state that reduces the efficacy of chemotherapy and predisposes bacteria toward evolving permanent antibiotic resistance [2]. The mechanisms of persister formation are in question, but in many laboratory conditions starvation and the stringent response appear to be key correlates in *E. coli* [3]. We reported enriched persister populations of *E. coli* in a condition of metabolic toxicity brought about by excess lactose [1]. Either this phenomenon occurs through a paradoxical stringent response, or there is an alternative pathway to persistence. To compare these hypotheses and measure gene expression, we performed RNA-Seq in starving, moderate, and toxic lactose conditions with or without antibiotic.

II. Enrichment Analysis Reveals Pathway-Level Regulation in Different Growth Conditions

Differential gene expression analysis was performed with the R package *DESeq2* [4]; pathway and GO term enrichment analysis was performed with the R package *pathfindR* [5].

A. Gene expression in starvation and toxicity

With *DESeq2* we used a general linear function for normalizing the count data, and multiple shrinkage algorithms to counter batch effects. Setting the sequencing profile for cultures in moderate lactose concentrations as the reference and the false discovery rate threshold to 0.1, we found substantial differences in the differentially expressed genes between starving and toxified conditions in comparison to moderate conditions. Most differentially expressed genes vary monotonically in their expression as lactose concentrations are varied, but a small subset does not, including an antitoxin and the nitrite reductase operon.

B. Pathways to persistence

We used *pathfindR* to enrich active subnetworks that link significant genes via protein-protein interaction. We obtained 12 enriched pathways in the starvation condition, and 9 pathways in the toxicity condition. Five of the pathways are shared by the starving and toxified conditions but are regulated differently. In starvation, two-component system pathways are mostly upregulated, suggesting limited proliferation while maximizing stress resistance. In toxified conditions, some transcription and translation related pathways are upregulated while DNA replication is downregulated.

III. Conclusions

Following a differential gene expression workflow, we discovered that bacterial pathways are mostly, but not always, monotonically regulated. Our results specifically suggest that interactions between the cell cycle, energy generation, and metabolic regulation guide the formation of a complex persister phenotype that does not show signs of starvation. Complete cell regulatory networks are difficult to analyze, but our results may provide a spark for generating a generalized phenotype model based on pathway regulation.

References

[1] Ray JC, Wickersheim ML, *et al.* 2016 Cellular Growth Arrest and Persistence from Enzyme Saturation. PLoS Comput Biol. 12(3):e1004825.
[2] Levin-Reisman I, Ronin I, *et al.* 2017 Antibiotic tolerance facilitates the evolution of resistance. Science. 355(6327):826-30.
[3] Harms A, Fino C, *et al.* 2017 Prophages and Growth Dynamics Confound Experimental Results with Antibiotic-Tolerant Persister Cells. MBio. 8(6).
[4] Love MI, Huber W, *et al.* 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15(12):550.
[5] Ulgen E, Ozisik O, *et al.* 2018 pathfindR: An R Package for Pathway Enrichment Analysis Utilizing Active Subnetworks. bioRxiv.272450.

Acknowledgements: This work was funded by NIH grant P20GM103638

[1]Center for Computational Biology, University of Kansas, Lawrence.
[2]Department of Molecular Biosciences, University of Kansas, Lawrence.
[3]E-mail: h749w664@ku.edu or jjray@ku.edu

Massive computational acceleration by using neural networks to emulate mechanism-based biological models

Shangying Wang[1], Kai Fan[2], Nan Luo[1], Yangxiaolu Cao[1], Feilun Wu[1], Carolyn Zhang[1], Katherine A. Heller[2] and Lingchong You[1,3,4,+]

***Abstract* — Mechanism-based mathematical models are the foundation for diverse applications. It is often critical to explore the massive parametric space. However, for many applications, such as agent-based models, partial differential equations, and stochastic differential equations, this practice can impose a prohibitive computational demand. To overcome this limitation, we present a fundamentally new framework to improve computational efficiency by orders of magnitude. Our work can potentially be a platform for faster parametric space screening of biological models with user defined objectives.**

I. Introduction

IN the past several decades, mechanism-based model have become increasingly important for exploring dynamics of biological systems. Many applications of these models depend on extensive, numerical solutions of the underlying equations. When a model becomes more complex, e.g. by accounting for stochastic dynamics or spatial dynamics or by accounting for many reactions, large-scale numerical simulations can become computationally prohibitive. When this occurs, the use of a model is usually limited to an extremely small portion of the possible parametric space, which greatly limits the potential value of a model.

II. Results

We have developed a hybrid strategy to overcome this limitation. Our strategy entails the use of artificial neural network (NN) to emulate mechanism-based models to enable massive acceleration in the prediction of system outputs of interest. In particular, we use a mechanism-based model to generate a set of simulation results, which are used as the basis for training the neural network. The simulated data set is sufficiently large for accurate training, yet small enough to be generated in manageable time window. When sufficiently trained, the neural network is then used to explore a parameter space that is several orders of magnitude greater than what can be explored by the original mechanism-based model (Figure 1).

In building our platform, our strategy entails two key innovations: Firstly, the use of long-short term memory networks to predict continuous outputs generated by mechanistic models. Secondly, development of a "voting" procedure (by paralleled trained neural networks) to enable self-assessment of prediction reliability by the neural networks.

To examine the potential of our methodology, we have applied it to two well-defined, published models [1-3]. In both examples, we have demonstrated massive acceleration of predictions with high accuracy by the trained neural networks. In addition to capture the qualitative features readily found in the training set, our analysis shows that the neural network exhibits rudimentary "creativity": it reliably predicts qualitative features rare or not even found in the training set. This property indicates the feasibility of using the neural network as a surrogate of the mechanistic model for discovering system properties not readily accessible by the mechanistic model.

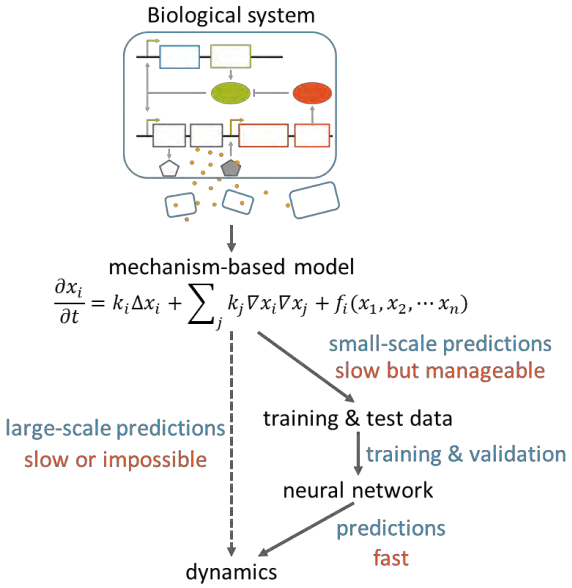


Figure 1 Using NNs to emulate a mechanism-based model

References

[1] Y. Cao, M. D. Ryser, S. Payne, B. Li, C. V. Rao, and L. You, "Collective Space-Sensing Coordinates Pattern Scaling in Engineered Bacteria," Cell, vol. 165, no. 3, pp. 620-30, Apr 21, 2016.

[2] G. Yao, T. J. Lee, S. Mori, J. R. Nevins, and L. C. You, "A bistable Rb-E2F switch underlies the restriction point," Nature Cell Biology, vol. 10, no. 4, pp. 476-U255, Apr, 2008.

[3] T. J. Lee, G. A. Yao, D. C. Bennett, J. R. Nevins, and L. C. You, "Stochastic E2F Activation and Reconciliation of Phenomenological Cell-Cycle Models," Plos Biology, vol. 8, no. 9, Sep, 2010.

Acknowledgements: This study was partially supported by the Office of Naval Research (N00014-12-1-0631), National Science Foundation (L.Y.) and National Institutes of Health (L.Y.: 1R01-GM098642; MDR: R01-GM096190).

[1]Department of Biomedical Engineering, Duke University, Durham, NC
[2]Department of Statistical Science, Duke University, Durham, NC
[3]Center for Genomic and Computational Biology, Duke University
[4]Molecular Genetics and Microbiology, Duke University, Durham, NC

Energy-dependent Regulation of the Cell Cycle

Shiyuan Wang[1], Qiong Yang[2]

***Short Abstract* — The cell cycle is driven by a mitotic circuit with an interlinked positive-plus-negative feedback loop topology. Although extensive research on the circuit's regulatory effect has been done, little is known about how cell's energy status is integrated into this circuit. By using the concentrated cycling extract of *Xenopus laevis* eggs, we investigated the effect of energy level on cell cycle oscillations. We found that while some aspects of the cell cycle remain unaffected, certain oscillatory characteristics, such as the period and amplitude, are sensitive to changes in ATP concentration. Furthermore, we observed an abrupt change of oscillatory behavior after continuous and prolonged cycles, which is presumably related to cell's prioritization caused by energy depletion.**

***Keywords* — ATP, energy, mitosis, biological oscillations, cycling extract, *Xenopus laevis*.**

I. Introduction

ATP fuels almost all energy-dependent biological processes, including the proliferation and division of individual cells. Although possible links were established between cell's energy status and mitosis [1][2], it is still debatable if and how ATP can affect the core mitotic circuit, which consists of the negative feedback loop between cyclin B-Cdk1 and APC/C and the interlinked positive and double-negative feedback loops involving the protein kinase Wee1 and phosphatase Cdc25 which collectively function as a bi-stable trigger [3]. To quantitatively manipulate the cell's ATP level and measure the cell's corresponding mitotic profile, we have developed a high-throughput cell-free system by encapsulating *Xenopus laevis* egg extracts into water-in-oil emulsion microdroplets [4][5]. These droplets function as artificial cells constantly undergoing mitotic oscillations. Due to its easily-accessible nature, the cell-free oscillatory system allows us to uncover important clock dynamics that would otherwise be difficult in intact cells.

II. Results

In order to visualize the cell cycle dynamics, we used a chimeric fluorescent protein, securin-CFP, which is a substrate of the APC/C. To study the effect of energy status on the cell cycle, we applied different amount of ATP to the *Xenopus* cycling extract as well as ATP live trackers such as QUEEN [6] to index every cell cycle with its energy level.

Preliminary results show that ATP concentration dramatically affects the cell cycle oscillation characteristics. Although the droplets have similar lifespans, droplets with higher ATP concentrations have significantly shorter periods and lower amplitudes. Oscillations in the same droplet also increases in period and amplitude over time. In addition, the period of securin degradation, which starts at anaphase and stops at the onset of interphase, does not change in respond to different ATP levels. The fluorescent track of securin-CFP also shows an increasing level of securin buildup through subsequent mitotic cycles, which might suggest a gradual positional shift of the ultrasensitive switch in the Cdk1–APC/C phase plane. One possible explanation is that the two positive feedback loops in the mitotic circuit involve the energy-consuming phosphorylation and dephosphorylation of wee1 and cdc25, and by changing the ATP concentration, we effectively change the strength of these two feedback loops, which in turn tune the mitotic clock to adjust to the current energy landscape to maintain its oscillatory status. However, further studies are required to verify the link.

Furthermore, we observed an abrupt change in oscillatory behavior in droplets with a prolonged lifespan (longer than ~25 hours). The period and amplitude experienced a sudden drop within one cycle. The exact nature of this change is still unclear, but it could be related to the cell shifting its priorities when sensing low-energy conditions.

References

[1] Koh, H., & Chung, J. (2007). AMPK links energy status to cell structure and mitosis. *Biochemical and Biophysical Research Communications*, 362(4), 789-792. doi:10.1016/j.bbrc.2007.08.016.

[2] Hardie, D., Scott, J. W., Pan, D. A., & Hudson, E. R. (2003). Management of cellular energy by the AMP-activated protein kinase system. *FEBS Letters*, 546(1), 113-120. doi:10.1016/s0014-5793(03)00560-x.

[3] Yang, Q., & Ferrell, J. E. (2013). The Cdk1–APC/C cell cycle oscillator circuit functions as a time-delayed, ultrasensitive switch. *Nature Cell Biology*, 15(5), 519-525. doi:10.1038/ncb2737.

[4] Guan, Y., Li, Z., Wang, S., Barnes, P. M., Liu, X., Xu, H., . . . Yang, Q. (2018). A robust and tunable mitotic oscillator in artificial cells. *ELife*, 7. doi:10.7554/elife.33549.

[5] Guan, Y., Wang, S., Jin, M., Xu, H., & Yang, Q. (2018). Reconstitution of Cell-cycle Oscillations in Microemulsions of Cell-free Xenopus Egg Extracts. *Journal of Visualized Experiments*, (139). doi:10.3791/58240

[6] Yaginuma, H., Kawai, S., Tabata, K. V., Tomiyama, K., Kakizuka, A., Komatsuzaki, T., . . . Imamura, H. (2014). Diversity in ATP concentrations in a single bacterial cell population revealed by quantitative single-cell imaging. *Scientific Reports*, 4(1). doi:10.1038/srep06522.

[1]Department of Biophysics, University of Michigan. E-mail: wangshi@umich.edu

[2]Department of Biophysics, University of Michigan . E-mail: qiongy@umich.edu

Inferring intercellular tension in cell-colony time-lapse videos

R. Vasan[1], M.M. Maleckar[2], C.D. Williams[2*], and P. Rangamani[1*]

***Abstract*— Actomyosin forces, membrane tension, and cell-cell adhesion produce changes in the shape of cell-cell interfaces. These shape changes can be exploited to estimate the local tension in colonies of cells. Here we present a method, Dynamic Local Intercellular Tension Estimation (DLITE), that infers cell-cell forces across time. DLITE is an extension and improvement upon existing, single time point, methods. We treat cell-cell junctions as a mechanical network, segmentable into circular arcs, whose morphology is dictated by the evolution of local tension and intracellular pressure.**

***Index Terms*— Force Inference, Tension, ZO-1, Cell Colony**

I. Purpose

DIRECT measurement of intercellular forces in confluent colonies of cells is challenging, requiring intervention in a densely packed space. Conducting such measurements over time is yet more challenging as subsequent force dynamics are disrupted by most direct forms of measurement, e.g. atomic force microscopy (AFM), micro-pipette aspiration, and retraction following laser ablation [1], [2]. Inferring cell-cell forces from the morphology of contact points provides an alternative, indirect, means of measuring intercellular forces. While this inferential approach is well suited for application to time series, it has primarily been applied to single time points [3]. Here we develop an alternate problem formulation that provides a non-destructive means to infer intercellular forces in time-lapse imaging of cell colonies. We term this technique Dynamic Local Intercellular Tension Estimation, or DLITE.

II. Summary of approach

Our work requires 1) the representation of the cell colony as a graph of connected circular arcs, and 2) the solving of a tension-balance for the resulting spatial graph. We computationally validate DLITE against synthetic data generated by the energy minimization framework, SurfaceEvolver. As a subsequent test-dataset we use time-lapse images of human induced pluripotent cells with endogenously labeled tight junctions (cell line AICS-23 from the Allen Institute for Cell Science). Using this *in vitro* image set as a base, we compare robustness to pathological cases, robustness to digitization error, and solution stability between DLITE and existing single-time-point techniques applied serially.

A. Colony representation

We define a colony as a directed planar graph comprising cells (c), edges (e) and nodes (n). Forces exerted by the actomyosin cortex result in tangential stresses in the form of tension (t) along an edge. Cells resist deformation by means of a normal stress exerted as pressure (p) inside every cell. Along each edge, we assume that the interfacial tensions are constant and that the intracellular pressures are uniform within a cell. At the length scale of the whole cell, we ignored membrane bending and assume that edge tensions and cell pressures exclusively govern cell shape. We assume that the colony is in quasi-static equilibrium at any point in time.

B. Solution finding

A general force balance at every node in a colony can be written as $n_{\mathrm{res}} = |\sum_{i=1}^{e_n} t_i v_i|$ where n_{res} is the residual force at node n, e_n is the number of edges connected to node n, and v represents the local tangent unit vector of an edge connected to n, defined by edge geometry. We frame the problem as an iterative optimization to minimize n_{res} at all nodes by choice of edge tension (t_i). Thus at every time-point we solve $\min_t f(t) = \sum_{j=1}^{N} (n_{\mathrm{j,res}} + r)$ subject to a regularization term, $r = n_{\mathrm{j,res}} / \sum_{i=1}^{e_j} \mid t_i v_i \mid$, where e_j is the number of edges connected to node j.

Results, briefly, are that DLITE is robust to two forms of network connectivity changes common in digitized time-series data and yields smoother tension-over-time predictions than do serially applied single-time-point approaches.

III. Conclusion

Inferring forces in time-series images of cell colonies provides us a low-impact, if indirect, means of observing the effects of forces on colony dynamics. Our technique and software enables this study.

We thank Paul G. Allen, founder of the Allen Institute for Cell Science, for his vision, encouragement and support. This work was partially supported by ARO W911NF1610411 to P.R.

[1]Department of Mechanical and Aerospace Engineering, University of California San Diego, San Diego, CA, 92122

[2]Allen Institute for Cell Science, Seattle, WA, 98109

*Corresponding authors' e-mails: cdavew@alleninstitute.org, prangamani@ucsd.edu

References

[1] W. J. Polacheck and C. S. Chen, "Measuring cell-generated forces: a guide to the available tools," *Nature methods*, vol. 13, no. 5, p. 415, 2016.

[2] M. Smutny *et al.*, "Uv laser ablation to measure cell and tissue-generated forces in the zebrafish embryo in vivo and ex vivo," *Methods in molecular biology*, vol. 1189, pp. 219–35, 2015.

[3] G. W. Brodland *et al.*, "Cellfit: a cellular force-inference toolkit using curvilinear cell boundaries," *PloS one*, vol. 9, p. e99116, 2014.

Connecting mechanisms of cell proliferation to size-homeostasis statistics

Lisa Willis[1], Henrik Jonsson[2], KC Huang[1,3,4]

Short Abstract — **High-throughput imaging protocols are rapidly generating cell-size homeostasis measurements among diverse organisms. For example, apparent "adder" behavior, a zero correlation between birth size and added size, holds across a range of bacteria, while eukaryotes can deviate from apparent adder. To understand the mechanistic implications of size-homeostasis statistics, we studied a general model of cell proliferation. Our analyses revealed plausible proliferation scenarios that yet fail to regulate cell size, conditions that generate adder behavior between birth and division, and how changes to these conditions cause deviations from adder behaviour as observed among eukaryotes.**

Keywords — **Cell-size homeostasis, growth and cell cycle regulation, mechanistic models, phenomenological models.**

I. Introduction

One of the most fundamental questions in biology is the how cells maintain their size, which can be intimately tied to myriad processes such as cell-cycle progression and transcription. Single-cell lineage tracking and cell-cycle reporters have led to a rapid proliferation in cell-size homeostasis measurements in bacteria, yeast, mammalian cells, and plant cells. Among bacteria [1,2,3] and an archaeon [4], a common theme has emerged: cells apparently regulate their size according to an "adder" behavior whereby a fixed volume is added between birth and division. Among eukaryotes, budding yeast and mammalian cells can deviate from adder behavior over G1 and S/G2 while maintaining apparent adder behavior between birth and division [5,6,7]. By contrast, stem cells of *Arabidopsis thaliana* exhibit intermediate adder-sizer behaviour [8]. Presently, despite some elegant theoretical papers [9,10], there is little clarity as to the mechanistic implications of these differences.

II. Results

We developed a simple and general model to systematically analyse the size-homeostasis behaviors produced by different cell proliferation scenarios, and consequently the constraints placed on proliferation by size control requirements. The scenarios include cell cycle checkpoints, G1/S or G2/M, regulated by master regulators (CDK1-cyclin in eukaryotes [11] or FtsZ/DnaA in bacteria [1,9]) that accumulate to critical activity levels, and inhibitor dilutors (Whi5 in budding yeast [5]) that are diluted-out by growth to trigger checkpoint progression at a minimum critical concentration, in combination with different regulator production, growth, division, and noise patterns. The main findings are:

A. G1/S inhibitor dilutors fail to achieve size homeostasis if the inhibitor is produced in proportion to cell size and S/G/M undergoes timer regulation, or if G1 duration is long (> 40% of the cell cycle) and G2/M undergoes critical size regulation.

B. Master regulators can lose cell-size control when production is gene copy-number limited or a critical concentration rather than a critical level triggers phase progression.

C. Inhibitor dilutors achieve adder-like, sizer-like, or supra-adder size homeostasis behaviors dependent on noise sources [10].

D. Apparent adder regulation between birth and division with non-adder regulation over G1 and S/G2/M arises naturally if G1/S or G2/M are independently regulated.

E. The intermediate adder-sizer behavior of *A. thaliana* stem cells can be attributed to alternative mechanisms with polarizing predictions for other size-homeostasis statistics.

References

[1] Willis, L., and Huang, K. (2017). Nat Rev Microbiol *15*, 606–620.
[2] Campos, M., Surovtsev, I.V., Kato, S., Paintdakhi, A., Beltran, B., Ebmeier, S.E., and Jacobs-Wagner, C. (2014). Cell *159*, 1433-1446.
[3] Taheri-Araghi, S., Bradde, S., Sauls, J.T., Hill, N.S., Levin, P.A., Paulsson, J., Vergassola, M., and Jun, S. (2015). Curr Biol *25*, 385-391.
[4] Eun, Y.-J., Ho, P.-Y., Kim, M., LaRussa, S., Robert, L., Renner, L.D., Schmid, A., Garner, E., and Amir, A. (2018).
[5] Schmoller, K.M., Turner, J.J., Koivomagi, M., and Skotheim, J.M. (2015). Nature *526*, 268-272.
[6] Chandler-Brown, D., Schmoller, K.M., Winetraub, Y., and Skotheim, J.M. (2017). Curr Biol *27*, 2774—2783.
[7] Cadart, C., Monnier, S., Grilli, J., Attia, R., Terriac, E., Baum, B., Cosentino-Lagomarsino, M., and Piel, M. (2018). Nat Commun *9*, 3275.
[8] Willis, L., Refahi, Y., Wightman, R., Landrein, B., Teles, J., Huang, K.C., Meyerowitz, E.M., and Jönsson, H. (2016). Proc Nat Acad Sci *113*, E8238—8246.
[9] Amir, A. (2014). Phys Rev Lett *112*, 208102.
[10] Barber, F., Ho, P.Y., Murray, A.W., and Amir, A. (2017). Front Cell Dev Biol *5*.
[11] Coudreuse, D., and Nurse, P. (2010). Nature *468*, 1074-1079.

[1]Department of Bioengineering, Stanford University, Stanford, CA, USA
[2]Sainsbury Laboratory, Cambridge University, Cambridge, UK
[3]Department of Microbiology and Immunology, Stanford University, Stanford, CA, USA
[4]Chan Zuckerberg Biohub, San Francisco, CA, USA

Bacterial Transport via Surface Tension Driven Flow

Ye Li[1], Liang Yang[2], and Yilin Wu[3]

***Short Abstract*— Most bacterial dispersal strategies rely on the motility of single cell. In contrast, most multicellular organisms rely on long-range directed fluid transport to maintain the physiological functions. We found prokaryote also can build complex large-scale transport system. We report a novel mechanism of long-range directed active transport in *Pseudomonas aeruginosa*, which does not rely on single-cell motility but requires surface-tension driven flows maintained by cooperative synthesis of bio-surfactants.**

***Keywords* — Long-range transport, cooperative behavior, surfactant.**

BACKGROUND

Long-range directed fluid transport is essential to maintain the physiological functions; it helps an organism to transport nutrients, metabolic wastes, and signaling molecules. By contrast, material transport in microbial world is often short-ranged and limited by diffusion (either passive diffusion due to thermal energy or active diffusion due to self-propulsion of motile cells [1-3]). Like in animals and plants, any form of long-range directed material transport would undoubtedly bring profound effect to the development, structure, and stress response of bacterial communities. Intriguingly, several examples of long-range directed transport in bacterial communities were reported in recent years [4-7]. These reported examples of directed material transport in bacterial communities were still limited in length scale, magnitude and controllability when compared to those in animals and plants.

RESULTS

Here we combined experiment and computational modeling to investigate a new form of bacterial transport behavior in prokaryotes powered by surface tension gradient. We found a potential motion pattern which does not rely on single-cell motility but requires surface-tension driven flows maintained by cooperative synthesis of bio-surfactants.

CONCLUSION

Our findings shed light on the cooperative behavior and long-range transport of bacteria, and provide new insight to the dynamics and function of bio-surfactants in bacterial communities.

REFERENCES

[1] Wu, X. L., & Libchaber, A. (2000). Particle diffusion in a quasi-two-dimensional bacterial bath. *Physical review letters*, *84*(13), 3017.

[2] Sokolov, A., Goldstein, R. E., Feldchtein, F. I., & Aranson, I. S. (2009). Enhanced mixing and spatial instability in concentrated bacterial suspensions. *Physical Review E*, *80*(3), 031903.

[3] Wu, Y., & Berg, H. C. (2012). Water reservoir maintained by cell growth fuels the spreading of a bacterial swarm. Proceedings of the National Academy of Sciences, 109(11), 4128-4133.

[4] Wu, Y., Hosu, B. G., & Berg, H. C. (2011). Microbubbles reveal chiral fluid flows in bacterial swarms. Proceedings of the National Academy of Sciences, 108(10), 4147-4151.

[5] Ping, L., Wu, Y., Hosu, B. G., Tang, J. X., & Berg, H. C. (2014). Osmotic pressure in a bacterial swarm. *Biophysical journal*, *107*(4), 871-878.

[6] Wilking, J. N., Zaburdaev, V., De Volder, M., Losick, R., Brenner, M. P., & Weitz, D. A. (2013). Liquid transport facilitated by channels in Bacillus subtilis biofilms. *Proceedings of the National Academy of Sciences*, *110*(3), 848-852.

[7] Wilking, J. N., Zaburdaev, V., De Volder, M., Losick, R., Brenner, M. P., & Weitz, D. A. (2013). Liquid transport facilitated by channels in Bacillus subtilis biofilms. *Proceedings of the National Academy of Sciences*, *110*(3), 848-852.

Acknowledgements: This work was funded by RGC grant 2130439, 2130493 and NSFC grant 21473152.

[1] Department of Physics, The Chinese University of Hong Kong, Hong Kong SAR, China. E-mail: skyliustc@gmail.com

[2] School of Medicine, Southern University of Science and Technology, Shenzhen, China. E-mail: yangl@sustc.edu.cn

[3] Department of Physics and Shenzhen Research Institute, The Chinese University of Hong Kong, Hong Kong SAR, China. E-mail: ylwu@cuhk.edu.hk

Modeling Complex Ecosystem in Collective Cancer Invasion Using an Evolutionary Game Theory Framework

Bin Zhang[1], Adam Marcus[2] and Yi Jiang[1]

***Short Abstract* — Cancer is an evolutionary disease [1] which exhibits genomic and phenotypic heterogeneity. Within a tumor, the subclones together with complicated surrounding micro-environments form complex multi-cellular ecosystems [2]. Tumors comprise a variety of specialized phenotypical subclones adapted to different ecological conditions that influence the response to treatments and prognosis of the diseases [3]. Recent experiments revealed existence of distinct phenotypes of cancer cells, leaders and followers, in non-small cell lung cancer during collective invasion [4]. We adopt an evolutionary game theory framework [5] to model the cancer microenvironments and the interactions between different phenotypes and quantify the proportion of phenotypes within a tumor. Our model reveals the pairwise interactions between leader and follower cells could alter the collective dynamics and suggests new treatment strategies within the tumor ecosystem.**

***Keywords* — Evolutionary Game Theory, Intratumor Heterogeneity, Collective Invasion, Leader-Follower**

I. INTRODUCTION

Non small cell lung cancer (NSCLC) constitutes 80% to 85% of lung cancers causing 120,000 deaths per year in the United States [6]. One of the main barriers for successful treatment is the tumor heterogeneity due to evolution [1] (i.e. chemotherapy sensitivity, metastatic potential and etc.). These subclones together with complicated surrounding micro-environments form complex multi-cellular ecosystem [2]. The evolution of the tumor and environmental selection alter the tumor dynamics and could have massive influences on tumor growth and its response to the treatments.

Using an image-guided genomics approach (SaGA), we confirmed the existence of at least two distinct phenotypes in NSCLC invasion packs: highly migratory leader cells and highly proliferative follower cells [2, 4]. In this study, we build an evolutionary game theory model [3, 5] to investigate the collective invasion dynamics on cell population level.

II. METHODS

Our model constructs a payoff matrix considering the pairwise interaction between leader cells (L) and follower cells (F). The intrinsic proliferation rate differences between the leaders and the followers describe their phenotypical fitness difference. Competition within the leader and follower cell population due to the limitation of the resources; interactions between the leaders and the followers and the potential benefits gained because of collective invasion all have major effects on the outcome of the collective dynamics.

A. Leader-Follower Composition

The intensity of the competition, interaction and potential gained benefits could shape the leader-follower composition which alters the collective dynamics. The observed 5% leader for optimal collective invasion can be explained using the payoff matrix.

B. Chemotherapy Schedule and Potential New Targets

The proportional balance between different tumor phenotypes can dramatically change with treatment conditions. The treatment intensity and schedule could be studied under this framework. The experiments revealed the interactions between L and F are controlled by signals [2, 4], which suggests potential new treatment targets. The model is an ideal theoretical test bed to quantify the effects of the new targeted chemotherapy agents.

III. CONCLUSION

Our model reveals that the pairwise interactions between leader and follower cells could alter the collective dynamics. The distinct interactions terms in the payoff matrix originated from leader-follower signals may suggests new drug targets for the treatments which reduces tumor burden as well as lower the metastatic risk.

REFERENCES

[1] Merlo, Lauren MF, et al. "Cancer as an evolutionary and ecological process." Nature reviews cancer 6.12 (2006): 924.
[2] Haney, Seth, et al. "The complex ecosystem in non small cell lung cancer invasion." PLoS computational biology 14.5 (2018): e1006131.
[3] Böttcher, Marvin A., et al. "Modeling treatment-dependent glioma growth including a dormant tumor cell subpopulation." BMC cancer 18.1 (2018): 376.
[4] Konen, J., et al. "Image-guided genomics of phenotypically heterogeneous populations reveals vascular signalling during symbiotic collective cancer invasion." Nature communications8 (2017): 15078.
[5] S Hummert, Sabine, et al. "Evolutionary game theory: cells as players." Molecular BioSystems 10.12 (2014): 3044-3065.
[6] American Cancer Society. Cancer Facts & Figures 2016.

[1]Department of Mathematics and Statsitics, Georgia State University. E-mail: bzhang17@student.gsu.edu & yjiang12@gsu.edu
[2]Department of Hematology & Medical Oncology, Winship Cancer Institute of Emory University. E-mail: admarcu@emory.edu

Understanding the Timing and System Architecture of the Mesenchymal Stem Cell to Bone-fat Two-way Fate Decision

Zhibo Zhang[1], Ting-Huan Chen[1], and Mary N. Teruel[1]

***Short Abstract* — Bone and fat cells are derived from a common precursor cell, mesenchymal stem cells (MSCs), and the imbalance of MSCs to these cell fates contributes to reduced bone mass and increased marrow adipose tissue in osteoporosis and skeletal aging. However, there is still a rudimentary understanding of the developmental genetic architecture controlling these cell-fate decisions. Here, we are developing unique genetic tools to study the dynamics of MSC lineage commitment at a single-cell level. Combined with mathematical modelling, these tools will allow us to understand the molecular basis of decision processes controlling MSC to adipocyte or osteoblast.**

***Keywords* — Osteogenesis, adipogenesis, cell fate decision, signaling dynamics.**

I. Background

THE concept of epigenetic landscape was proposed by Conrad Hal Waddington to illustrate the pattern of cell dynamics during development. It shows that the multipotent cell goes through a series of binary branching points to differentiate successively into distinct cell types [1]. The gene circuits that govern fate decisions are characterized by the mutual inhibition of two opposing fate-determining transcription factors (TFs), which each repress the activity of the other through a variety of molecular mechanisms, typically involving protein–protein interactions. Cross-inhibition produces bistability with two attractors, which represent two differentiated cell fates respectively [2]. However, we found that, during the osteogenesis of C3H10T1/2 pluripotent stem cells, the mRNA level of PPARG, the master TF in the later stage of adipogenesis [3], was remarkably increased in a similar manner of that of Osterix, the master TF of bone-forming cells [4]. In addition, both PPARG and Osterix protein showed significant nuclear localization above background under osteogenic stimuli. The synchronized expression pattern of PPARG and Osterix indicated that there might be a novel regulatory mechanism underlying bone-fat fate decision.

II. Main hypothesis

It has been reported that Osterix repressed adipogenesis by negatively regulating PPARG transcriptional activity. Meanwhile, Osterix was upregulated during the later stages of adipogenesis [5]. Thus, we supposed that PPARG and Osterix constitute a negative-feedback loop. Theoretically, negative feedback is not able to generate binary attractors. But using mathematical modeling, we showed that negative feedback loop can regulate the relative timing of two TFs via decelerating the accumulation rate of the positive TF (i.e. PPARG) [6]. The model indicated that it was the differentiation paths in phase space that determined the cell lineage, rather than the positions of attractors. It's potential that the stem cell makes decision to choose lineage by detecting which TF is the first to be activated.

To verify our hypothesis, we are developing live single-cell imaging platform to simultaneously monitor the dynamics of three key TFs (i.e. Runx2, PPARG and Osterix) under various differentiation signals and systematical genetic perturbations. These experiments will provide a much clearer picture about how cells encode their extracellular signals into the relative timing between interlinked regulators, and how down-stream genes decode timing information to make a proper fate decision.

III. Conclusion

This work will reveal the interactions between Osterix and PPARG, and that how the timing order of these two fate regulators is regulated to control lineage commitment of MSCs to adipocytes and osteoblasts. This new mechanism will help develop timing-based strategies to artificially switch cell fates, and to identify new targets for the treatment of bone diseases such as osteoporosis.

References

[1] Waddington CH (1957) *The strategy of the genes: a discussion of some aspects of theoretical biology.* Allen & Unwin, London.

[2] Zhou JX & Huang S (2011) Understanding gene circuits at cell-fate branch points for rational cell reprogramming. *Trends in Genetics* **27.2**, 55-62.

[3] Brun RP (1996) Differential activation of adipogenesis by multiple PPAR isoforms. *Genes & Development* **10.8**, 974-984.

[4] Nakashima K (2002) The novel zinc finger-containing transcription factor osterix is required for osteoblast differentiation and bone formation. *Cell* **108**, 17-29.

[5] Han Y, et al. (2016) Osterix represses adipogenesis by negatively regulating PPARγ transcriptional activity. *Scientific Reports* **6**, 35655.

[6] Bahrami-Nejad Z et al (2018) A Transcriptional Circuit Filters Oscillating Circadian Hormonal Inputs to Regulate Fat Cell Differentiation. *Cell Metabolism* **27.4**, 854-868.e8.

[1]Department of Chemical and Systems Biology, Stanford University, Stanford, CA 94305, USA

Binary establishment and maintainance of discrete cell fates in development

Jiaxi Zhao[1,3], Jacques Bothma[2], Matthew Norstad[2] and Hernan G. Garcia[1,2,4]

***Abstract*— During embryonic development, cells must ultimately adopt discrete fates. Positive autoregulation has been proposed as a general mechanism for establishing and maintaining these discrete cell-fate decisions. Here, we quantitatively dissect the role of autoregulation and bistability in establishing binary cellular fates in the fruit fly *Drosophila melanogaster*. Specifically, we apply recently developed single-cell live imaging techniques to quantify transcriptional and protein dynamics of the *Drosophila* pair-rule gene *fushi tarazu* as cells decide whether to commit to the expression of the gene.**

***Index Terms*— Autoregulation, *Drosophila* embryonic development, Transcription dynamics, Protein dynamics**

I. Introduction

ONE of the mysteries of animal development is how the precise control of gene expression and developmental fates is achieved. This ranges from the rapid control of gene expression dynamics to maintaining constant protein levels throughout the developmental process to ensure the stable and irreversible adoption of cellular fates. Autoregulation is one of the key network motifs speculated to "lock" transcription patterns and maintain gene expression levels at precise values which are essential for the correct establishment of developmental fates. During embryogenesis, cells need to make binary decisions through the all or none "switch-like" expression of genes. It has been proposed that bistability arising from positive autoregulation contributes to this switch-like behavior.

Although this relatively simple mechanism has been studied in the context of single-cell organisms, including bacteria and budding yeast, it has never been quantitatively explored in the context of a developing multicellular organism.

II. Model System

The *Drosophila* pair-rule gene *fushi tarazu (ftz)* is expressed as a seven-stripe pattern in the early embryo [1]. It has been shown that two important enhancer elements dictate *ftz* gene expression dynamics during early embryogenesis. The early element responds to upstream transcription factors in the gene regulatory network and sets the initial expression pattern of seven stripes. The late element sharpens and maintains the pattern through its autoactivation by Ftz protein. [2].

[1]Department of Physics, University of California, Berkeley.

[2]Department of Molecular and Cell Biology, University of California, Berkeley.

[3]Email: jiaxi.zhao@berkeley.edu

[4]Email: hggarcia@berkeley.edu

III. Method

We are applying recently developed techniques to quantitatively measure transcriptional dynamics and protein dynamics of *ftz* autoregulation in single cells of living fruit fly embryos [3][4].

Our objective is to generate a phase diagram that reveals the dynamics and quantitative basis of this switch-like decision. To make this possible, we have "separated" the pattern in order to simultaneously measure the transcription and protein dynamics dictated by the early and late regulatory elements.

Specifically, we are using the EGFP-LlamaTag system to measure temporal protein dynamics of the early element. In addition, transcriptional output of the *ftz* autoregulatory element will be measured with a combination of the MCP-MS2 and mCherry-LlamaTag systems.

With this capability to measure the dynamics of autoregulation in a developing embryo, we will test a simple theoretical model of this switch-like system by changing *ftz* mRNA or protein half-life based on known mutations [5][6].

IV. Expected Results

With our capability for measuring and perturbing the autoregulation process, we will explore:

1) Whether developmental commitment can be understood as a simple bistable dynamical system.
2) The contribution of diffusion and transcriptional bursting to developmental commitment.

By exploring the above questions, this study will shed light on the quantitative principles by which binary decisions are made in development.

References

[1] Y. Hiromi, A. Kuroiwa, and W. J. Gehring, "Control elements of the Drosophila segmentation gene *fushi tarazu*," *Cell*, vol. 43, no. 3, pp. 603–613, 1985.

[2] Y. Hiromi and W. J. Gehring, "Regulation and function of the Drosophila segmentation gene *fushi tarazu*," *Cell*, vol. 50, no. 6, pp. 963–974, 1987.

[3] J. P. Bothma, M. R. Norstad, S. Alamos, and H. G. Garcia, "LlamaTags: a versatile tool to image transcription factor dynamics in live embryos," *Cell*, vol. 173, no. 7, pp. 1810–1822, 2018.

[4] H. G. Garcia, M. Tikhonov, A. Lin, and T. Gregor, "Quantitative imaging of transcription in living Drosophila embryos links polymerase activity to patterning," *Current Biology*, vol. 23, no. 21, pp. 2140–2145, 2013.

[5] K. Kellerman, D. M. Mattson, and I. Duncan, "Mutations affecting the stability of the *fushi tarazu* protein of Drosophila." *Genes & Development*, vol. 4, no. 11, pp. 1936–1950, 1990.

[6] A. Riedl and M. Jacobs-Lorena, "Determinants of Drosophila *fushi tarazu* mRNA instability." *Molecular and cellular biology*, vol. 16, no. 6, pp. 3047–3053, 1996.