

Emerging Mathematical Challenges in Synthetic Biological Network Design

Enoch Yeung (University of California, Santa Barbara),
Andras Gyorgy (New York University Abu Dhabi),
Mustafa Khammash (ETH Zurich),
Ophelia Venturelli (University of Wisconsin-Madison)

August 25–27, 2023

Thirty one researchers met in hybrid format at the Banff International Research Station to discuss emerging mathematical challenges in designing synthetic gene networks. The workshop was organized around the following themes: (i) distributed & multi-cellular biological control; (ii) from modularity to robustness; (iii) biological context & control; and (iv) quantitative design & discovery.

1 Overview of the Field

Synthetic biology seeks to understand and control living processes by building them. Nearly twenty years after its inception, the field now faces emerging challenges as it builds increasingly complex genetic circuits and networks. These challenges can be categorized roughly into five areas: (i) scalability, (ii) robustness, (iii) modularity, (iv) context, and (v) data-driven approaches.

First, genetic circuits rely on biochemical energy borrowed from living cells. As a result, seemingly disconnected network components influence each other due to the limited availability of shared cellular resources [5, 11]. Regarding the second challenge, the cellular environment is replete with uncertainty due to the lack of precise knowledge of endogenous and engineered molecular components and their interactions, in addition to external perturbations and stochastic noise. As a result, models of cellular function are necessarily uncertain, and designing circuits that function reliably in the face of this uncertainty requires circuit architectures that are robust by design [12]. The third major challenge that synthetic biologists often face centers on modularity. The ability to accurately predict the behavior of a complex system from that of the composing modules has been instrumental to the development of engineering systems. Unfortunately, despite the fact that biological networks are rich of frequently repeated motifs [1], suggesting a modular organization, our current ability of predicting the emergent behavior of a network from that of the composing modules remains limited. Fourth, while synthetic biology bears many similarities to traditional engineering disciplines, designing synthetic gene circuits is often time consuming due to their context-dependent behavior, frequently leading to unexpected and perplexing phenomena [4]. Thus, the construction of even simple systems typically relies on massive DNA libraries that need to be iteratively refined, involving high-throughput screening and testing in a lengthy and expensive process [13, 3]. Finally, synthetic biological networks are consistently represented through experimental data, but translating data to generalizable biophysical insight and parameters remains a challenge [8]. While multiple data streams have become commonplace in characterizing engineered living systems, data-driven methods for extracting and embedding insight from these into an integrated model for learning representation are lacking.

2 Recent Developments and Open Problems

Recent years have seen tremendous progress in synthetic biology ranging from the development and implementation of insulation devices to multiple controllers facilitating both the modular and rational design as well as the robust functioning of genetic circuits [6, 15, 2, 10, 7, 9]. These results have been made possible by combining mathematical tools from a variety of fields with experimental techniques.

For instance, in addition to distributing complex genetic functions across multiple heterogeneously programmed cells [14], mathematical tools aided by experimental approaches were also developed to capture the previously unmodeled interactions among circuit components to overcome the challenge of scalability [17]. Motivated by the success of this approach, there is an urgent need for developing new mathematical (and experimental) approaches for engineering synthetic microbiomes and consortia [16]. While control theory from engineering offers some clues on the structural requirements of robust designs [12], the unique nature of molecular systems necessitates the development of new mathematical methods and accompanying experimental tools for characterizing and engineering biological circuits that are insensitive to the highly uncertain environment peculiar to the living cell. Additionally, it has been proposed that biological networks may have a modular organization similar to that of engineered systems, and that core processes, or motifs, have been conserved through the course of evolution and across different contexts [1]. In addition to having profound consequences from an evolutionary perspective, this view implies that biology can be understood, just like engineering, in a modular fashion. Unfortunately, genetic circuits often display context-dependent behaviors, undermining the modular design of complex systems [5, 11]. Therefore, it is imperative to properly model and abstract biological interconnections and the emergent behavior of interconnected, low-dimensional, non-linear systems upon their composition. While the iterative re-design and tuning of synthetic systems can prove successful for modules of modest complexity [13, 3], it rapidly becomes infeasible with increasing circuit size. Context-dependence thus poses a critical limitation in synthetic biology by undermining the rational design of large-scale systems [4]. It is becoming clear that new mathematical and experimental approaches are urgently needed to facilitate a deeper understanding of context-dependence and to devise practical approaches for overcoming it. Finally, despite the overabundance of biological data (e.g., transcriptomics, proteomics, flow cytometry, time-lapse single cell microscopy, photospectrometry), a generalized mathematical theory integrating experimental measurements with biological system representation is lacking for the development of data-driven approaches for engineering complex synthetic biological systems [8].

3 Presentation Highlights

The focus of this workshop was the coordination, unification, and cross-fertilization of myriad mathematical methods in designing synthetic biological gene networks. The informal yet focused environment provided by the BIRS station was the perfect place to have critical conversations about technology gaps and how to increase scalability, robustness, and define the next generation of challenge problems in biological design. This has been highlighted by some of the testimonials: *This BIRS workshop was very stimulating and interesting: I could hear about new results, meet friends as well as new contacts (whom I would have probably never met at the conferences that I typically attend) and learn a lot, especially thanks to the interdisciplinary nature of the workshop, and have new research ideas (Giulia Giordano); The BIRS workshop on “Emerging Mathematical Challenges in Synthetic Biological Network Design” provided a unique venue for like-minded synthetic biologists to discuss their most recent research results in excellent mathematical detail. Unlike other meetings in this field, presenters could provide complete derivations and explanations of their results without fear of losing their audience (Howard Salis).* The hybrid format allowed us to have not only detailed discussion in person, but also accommodate the schedule of those who could not attend personally.

4 Scientific Progress Made

To address the major challenges highlighted above, progress has been made in four thematic areas: (i) distributed and multi-cellular biological control; (ii) modularity and robustness; (iii) biological context and control; and (iv) quantitative design and discovery.

First, considering distributed and multi-cellular biological control, we learned about the possibility of predicting and controlling gene transfer in microbial communities (Lingchong You) and their heterogeneity due to stochastic promoter switching (Yili Qian), as well as how to implement robust dynamical biosystems with layered controls (Chelsea Hu) and the principles of microbial community efficiency, robustness, and controllability (Ophelia Venturelli). These experimental talks were complemented by theoretical results ranging from reaction order analysis to reveal global polyhedral constraints on the behavior of biomolecular reaction systems (John Marken) to Bayesian optimization of microbiomes using tailored machine learning (Jaron Thompson).

Second, participants highlighted the role that modularity plays to facilitate achieving robustness. For instance, we learned not only about the structural and topology-independent stability of biological systems (Giulia Giordano) but also how to characterize and implement maximally robust genetic control circuits (Mustafa Khammash). In addition to quantifying the impact of resource limitations (Eduardo Sontag), a general framework was presented that quantifies the contribution of a biological feedback control mechanism to adaptation (Mariana Gmez-Schiavon), as well as concrete application examples of biological control for the automatic detection of gut infection (Michaëlle Mayalu) and for the governing of bottom-up engineered bacteria consortia by synthetic gene circuits (Xiao Wang).

Third, it was illustrated not only how supercoiling can couple the behavior of unrelated genes (Enoch Yeung), but also how inducible plasmid copy number control can facilitate the characterization of these surprising interactions (Andras Gyorgy) with the help of the proper dynamical framework for characterizing gene expression dynamics (Terence Hwa). Furthermore, cell-free transcription-translation systems were highlighted as an efficient platform to unentangle the complicated relationship of genetic circuits and their context (Vincent Noireaux) and ribocomputing as a promising tool for implementing large-scale cellular computations (Jongmin Kim).

Fourth, illustrating the role of quantitative design and discovery, talks centered around the mutually beneficial fusion of computational approaches and experimental techniques. Topics ranged from the rational design of artificial phosphorylation signaling networks in human cells (Caleb Bashor) to engineering bacteria for mastering tic-tac-toe through accelerated adaptation (Alfonso Jaramillo) and to the control of complex biological systems interfaced with bioelectronics, showcasing the current frontier of experimental approaches. These were complemented by cutting-edge computational techniques, including leveraging Fisher information to harvest fluctuation information while rejecting measurement noise (Brian Munsky), verification- and model-guided gene circuit design and prototyping (Lucia Marucci and Chris Myers), as well as the dynamic modelling of mixed microbial populations (Brian Ingalls).

5 Outcome of the Meeting

This workshop had four outcomes. First, presenting the latest advances from expert synthetic biologists to illustrate and highlight emerging challenges in biological scale-up, modular design, context robustness, data-driven representation. Second, allowing theorists and engineers to share their recent advances in mathematical modeling and data-driven design that can aid in addressing these challenges. Third, incorporating the input from the biologists to spur development and organization of new mathematical strategies or approaches, integrating ideas from control theory, systems theory, nonlinear analysis, functional analysis, behavioral systems theory, and pure mathematics to address the fundamental differences in these emerging challenges. Fourth, providing a venue for theoreticians to forge new collaborations with leading practitioners in systems and synthetic biology.

Invited speakers were selected to ensure accessibility for a mixed audience of experimentalists and theorists. Focused discussions on challenges of interest emphasized the need to solicit multiple perspectives from practitioners and theorists. Key junior speakers were identified who serve as emerging thought leaders in synthetic biology and biological control, specifically those who represented underrepresented minorities in the STEM fields and our field at large. We also placed an emphasis on inviting more female engineers and scientists, which historically have been underrepresented in the field of synthetic biology and biological control. Additionally, the international roster of speakers represented great geographical diversity from North and Central America, Europe, Asia, and the MENA region.

References

- [1] Uri Alon. Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8(6):450–461, June 2007.
- [2] Stanislav Anastassov, Maurice Filo, Ching-Hsiang Chang, and Mustafa Khammash. A cybergenetic framework for engineering intein-mediated integral feedback control systems. *Nature Communications*, 14(1):1337, March 2023.
- [3] Evan Appleton, Curtis Madsen, Nicholas Roehner, and Douglas Densmore. Design Automation in Synthetic Biology. *Cold Spring Harbor Perspectives in Biology*, 9(4):a023978, April 2017.
- [4] Stefano Cardinale and Adam Paul Arkin. Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems. *Biotechnology Journal*, 7(7):856–866, July 2012.
- [5] Francesca Ceroni, Rhys Algar, Guy-Bart Stan, and Tom Ellis. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nature Methods*, 12(5):415–418, May 2015.
- [6] Domitilla Del Vecchio, Alexander J Ninfa, and Eduardo D Sontag. Modular cell biology: retroactivity and insulation. *Molecular Systems Biology*, 4(1):161, January 2008.
- [7] Maurice Filo, Sant Kumar, and Mustafa Khammash. A hierarchy of biomolecular proportional-integral-derivative feedback controllers for robust perfect adaptation and dynamic performance. *Nature Communications*, 13(1):2119, April 2022.
- [8] Paul S. Freemont. Synthetic biology industry: data-driven design is creating new opportunities in biotechnology. *Emerging Topics in Life Sciences*, 3(5):651–657, November 2019.
- [9] Timothy Frei, Federica Cella, Fabiana Tedeschi, Joaquin Gutierrez, Guy-Bart Stan, Mustafa Khammash, and Velia Siciliano. Characterization and mitigation of gene expression burden in mammalian cells. *Nature Communications*, 11(1):4641, September 2020.
- [10] Timothy Frei, Ching-Hsiang Chang, Maurice Filo, Asterios Arampatzis, and Mustafa Khammash. A genetic mammalian proportionalintegral feedback control circuit for robust and precise gene regulation. *Proceedings of the National Academy of Sciences*, 119(24):e2122132119, June 2022.
- [11] Andras Gyorgy, JosI. Jimnez, John Yazbek, Hsin-Ho Huang, Hattie Chung, Ron Weiss, and Domitilla DelVecchio. Isocost Lines Describe the Cellular Economy of Genetic Circuits. *Biophysical Journal*, 109(3):639–646, August 2015.
- [12] Mustafa H. Khammash. Perfect adaptation in biology. *Cell Systems*, 12(6):509–521, June 2021.
- [13] Kevin D Litcofsky, Raffi B Afeyan, Russell J Krom, Ahmad S Khalil, and James J Collins. Iterative plug-and-play methodology for constructing and modifying synthetic gene networks. *Nature Methods*, 9(11):1077–1080, November 2012.
- [14] Javier Maca, Francesc Posas, and Ricard V. Sol. Distributed computation: the new wave of synthetic biology devices. *Trends in Biotechnology*, 30(6):342–349, June 2012.
- [15] Deepak Mishra, Phillip M Rivera, Allen Lin, Domitilla Del Vecchio, and Ron Weiss. A load driver device for engineering modularity in biological networks. *Nature Biotechnology*, 32(12):1268–1275, December 2014.
- [16] Ophelia S. Venturelli, Harris H. Wang, Sylvie Estrela, Kerwyn Casey Huang, Rodrigo Ledesma-Amaro, Alex J.H. Fedorec, Brbel Stecher, Christopher E. Lawson, Amir Zarrinpar, Chun-Jun Guo, and Orkun S. Soyer. What is the key challenge in engineering microbiomes? *Cell Systems*, 14(2):85–90, February 2023.
- [17] Enoch Yeung, Aaron J. Dy, Kyle B. Martin, Andrew H. Ng, Domitilla Del Vecchio, James L. Beck, James J. Collins, and Richard M. Murray. Biophysical Constraints Arising from Compositional Context in Synthetic Gene Networks. *Cell Systems*, 5(1):11–24.e12, July 2017.