Squaring a circle: To what extent are traditional circuit analogies impeding synthetic biology?

Katherine Ilia[1,2]
Domitilla Del Vecchio[2,3]

[1]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.
[2]Synthetic Biology Center, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.
[3]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

Synthetic biology has been significantly shaped by modular design principles via analogies to electrical and computer engineering. While convenient, these parallels often break down in practice, and we are still largely unable to engineer sophisticated systems that behave as predicted. As nature has achieved robust and intricate programs without requiring strict modularity, we may want to revisit genetic circuit design approaches. Rather than pursuing modularity, we could aim for a robust and scalable design framework that embraces the uncertainty that context-dependence brings to engineering in a biological chassis. Systems and control theory offer a starting point, but a substantial conceptual leap will be needed to quantitatively predict system behavior and establish flexible context-aware design processes. Only by overcoming these hurdles, we will be able to capitalize on synthetic biology in particular and on biotechnology in general for medicine, environmental engineering, and energy production.

Auditing the analogies to electrical and computer engineering

While "synthetic biology" as coined in 1911 more so referred to synthetic chemistry than biology,[1,2] this term now represents a multidisciplinary field of science that involves redesigning biological processes and organisms for useful purposes. Its roots can be traced back to 1961, when Jacob and Monod introduced the operon model of gene regulation,[3–6] which was the core process used in 2000 to create the first two synthetic genetic networks.[7,8] Since then, "genetic circuit" has been used as a synonym for genetic network, as in earlier systems biology studies.[9,10] Shortly after the creation of these two synthetic genetic systems, some members of the engineering community advocated that synthetic biology adopt a similar design abstraction hierarchy as employed in electronic circuit design.[11,12] Within such an approach, design should be performed at different levels of abstraction (DNA, genetic parts, modules, and systems) and should allow one to disregard details of lower levels when designing at any given layer. Within any given layer, elements should then be composed through well-characterized physical or information interfaces to allow arbitrary combinations of elements into systems with predictable behavior.[11] Since then, these design principles have permeated the field as a basis for genetic circuit design.[11,13–16] In fact, abstracting genetic components as input/output (I/O) maps, often digital and static (as opposed to analog and dynamic), has been highly convenient for performing stepwise and systematic composition of complex circuits, as demonstrated by design tools such as Cello and COMET.[15,17]

However, for this approach to be applicable, defined genetic circuit components must be modular; that is, they must be fully characterizable by their I/O behavior and connectivity, wherein this behavior is maintained upon arbitrary composition (Figure 1, Left). By contrast, there is clear evidence that biological components, as defined and used today, do not satisfy this property

(Figure 1, Right).[18–45] Whereas electronic circuit components are often designed with internal compensation mechanisms to maintain pre-defined I/O properties independent of their context, today's genetic components are by no means guaranteed to be modular. Indeed, a component's behavior is contingent on its intracellular and extracellular context via diverse interactions, such as DNA supercoiling,[18] chromatin state,[19] positional effects,[20–23] off-target interactions,[24,25] retroactivity,[26,27] resource sharing,[28–33] cell fitness and competition,[34,35] microenvironmental cues,[36] random mutations,[37] and growth rate feedback,[38,39] to name a few. Most of these interactions are currently not in the description of genetic elements and of their connectivity. Furthermore, the biochemical reactions that drive circuit functionality are intrinsically stochastic,[40–43] I/O responses are more often analog than digital,[44] and temporal dynamics make static I/O characterization insufficient to capture emergent system behavior.[45]

To move the field forward, we need to balance the convenience and drawbacks of the assumptions implicit in a modular and hierarchical design approach. For biological engineers, these analogies may ultimately become an easily accessible starting point that aids early design stages. Since design is iterative, one could initially assume that each component achieves a well-defined subfunction with a fixed connectivity. For instance, one could assume that modules have strictly unidirectional interactions and that genetic parts, such as promoters, have the same activity independent of the surrounding DNA sequences. In subsequent design iterations, one could then model bidirectional flow of information among modules and additional physical interactions among parts, such as with retroactivity,[26] resource loading,[31] and DNA supercoiling.[18] In this framework, retroactivity, loading, parts-sharing, and interference may even become useful features, as opposed to bugs to be stamped out.

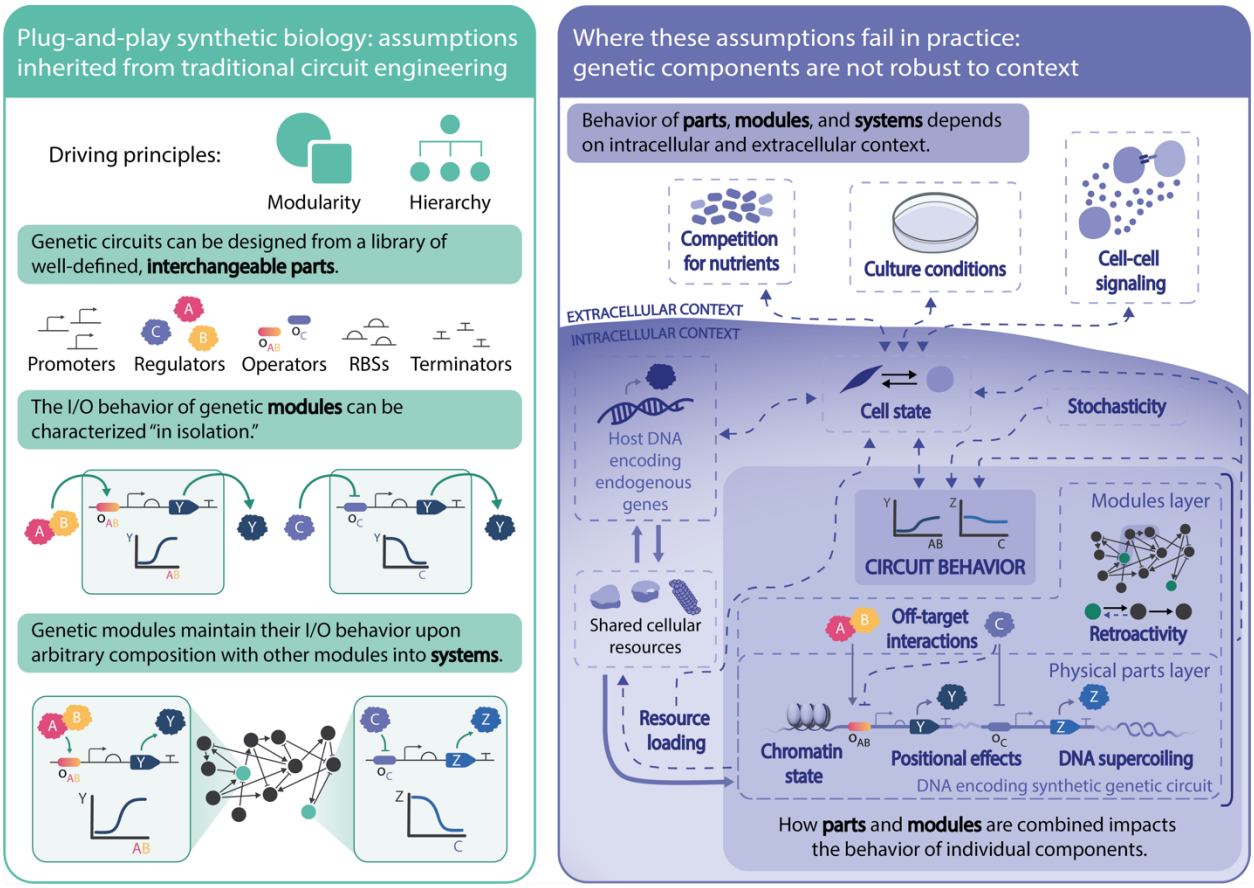

Figure 1: Auditing the analogies to electrical and computer engineering. (Left) The appeal of applying principles from traditional circuit engineering disciplines is in its simplicity. Use of these principles rests on assumptions inherent in modular and hierarchical design that simplify the design and optimization of genetic circuits. (Right) In practice, these assumptions often fail given the complexity of biological systems. For simplicity, many interactions among the factors illustrated here are only drawn indirectly via their role in modulating cell state. In this way, these factors become coupled and are difficult to disentangle. For instance, high transgene expression places a load on cellular resources, which causes endogenous gene expression to change, thereby potentially affecting growth rate (encapsulated here in cell state). In cases where there are multiple engineered cell types or strains in a shared culture, the cells in which growth rate is higher will tend to dominate the population. The dashed arrows highlight some of the interactions among factors that are often not accounted for in design and, depending on the application, may be undesirable. The set of solid arrows from "Shared cellular resources" to "Host DNA encoding endogenous genes" and "DNA encoding synthetic genetic circuit" indicate that cellular resources are required to produce both endogenous and synthetic gene products. At the same time, endogenous genes encode many cellular resources, such as the ribosomes, polymerase and proteosome depicted here, as illustrated by the solid arrow from "Host DNA encoding endogenous genes" to "Shared cellular resources." The dashed arrow from "DNA encoding synthetic genetic circuit" to "Shared cellular resources" indicates that cellular resources are loaded (temporarily sequestered) by the process of transgene expression. Note that chromatin state is primarily in eukaryotic cells, and RBSs are mainly used in bacterial synthetic biology. The term "positional effects" refers to any sequence-dependence of a part's performance, as can be the case for promoters and RBSs. Stochasticity includes processes that occur at low molecular counts and factors such as copy/integration number variation. See main text for relevant references. *Abbreviations: RBSs = ribosome binding sites.*

Rising to the unique challenges of building genetic circuits

The lack of strict modularity of engineered genetic components emanates, in part, from the poor robustness of a component's defined I/O behavior to context. In fact, lack of robustness to context

is a major barrier to modular and scalable design because existing elements in a system change properties once new components are added, and these elements consequently need to be redesigned. This way, design complexity increases combinatorially, rather than linearly, with the number of components in the system, thus rendering the design process monolithic and unscalable. Lack of robustness also curtails the practical impact of synthetic biology: most genetic circuits built today usually function as intended only in tightly controlled laboratory and cellular conditions (in general, reproducibility of research in biological sciences and systems biology modeling is low, as a result of a variety of factors, including lack of information reported on the experimental setup or modeling workflow[46–49]). With this fragility, it is difficult to envision a future where engineered organisms will be deployed in the field, whether for environmental biosensing or as therapeutic agents. Our current inability to efficiently and predictably design robust circuits stems from our limited understanding of how the properties of genetic parts, modules, and systems vary with genetic context,[18,50] intracellular conditions and connectivity,[28–30,33,38,51] and extracellular environment.[52,53] How can we tackle this formidable challenge? At the most basic level, we have two options: (a) insulating circuit components from their context so they behave as designed despite disturbances, or (b) evolving current modeling and design frameworks to enable prediction and optimization of complex interactions among modules and between a circuit and its context. To advance synthetic biology, we will likely need a combination of the two.

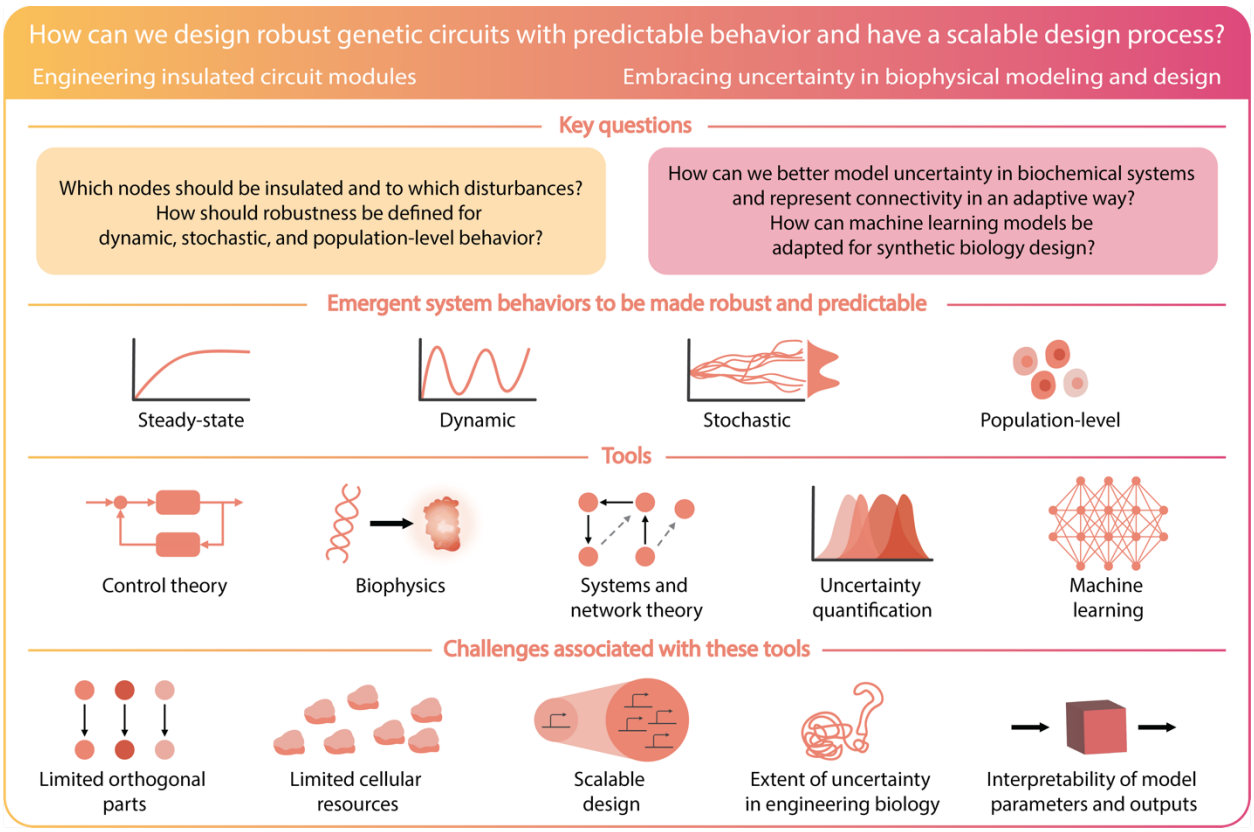

Figure 2: Rising to the unique challenges of building genetic circuits. To be able to design robust genetic circuits reliably and predictably, we will likely need a combination of two approaches: engineering insulated genetic modules and developing more sophisticated modeling and design frameworks that handle biological uncertainty (e.g., in connectivity among parts, simple physical process behavior, and parameters) and noise. Representative questions corresponding to engineering insulated circuit modules are shown in the yellow rectangle on the left, and those

corresponding to embracing uncertainty in biophysical modeling and design are shown in the pink rectangle on the right. Emergent system behaviors to be made robust and predictable, tools, and challenges associated with these tools are common to both approaches.

Designing insulated genetic circuit modules Robust circuit components are advantageous as they can maintain a desired behavior in the presence of environmental disturbances. There are already examples of genetic circuits that implement negative feedback and feed-forward compensation mechanisms to attenuate interference from select extra-modular processes.[29,30,54–64] As this work continues, we need to carefully consider which system properties should be robust and to which perturbations they should be robust. Additional challenges include the trade-offs between robustness to disturbances and input sensitivity, and between designing for robustness and maintaining a scalable design process. For instance, if we were to implement feedback compensation for every single module, we would quickly introduce even more resource loading and run out of orthogonal parts. This approach may also prevent us from exploiting additional inter-module interactions, thereby yielding suboptimal circuit designs. These practical challenges of engineering robust modules will require us to consider which nodes in a system require feedback compensation for ensuring robustness of emergent output properties. We will also need to develop new uncertainty representations and robustness techniques to address these and other aspects of engineering robust genetic circuits. In fact, while classical robust control methodologies[65] may be leveraged, these are mainly focused on exact parameterization of uncertainty, which is most likely not conducive for engineering biology. Therefore, new methods that tolerate large and unstructured uncertainty, also and especially within the controller components, will be needed. In addition, robustness is often only addressed for steady state or deterministic I/O behavior. However, it is critical that we define metrics for assessing robustness of a broader set of dynamical and stochastic properties, such as with respect to periodic behavior, multi-stability, probability distributions, and more sophisticated signal processing.

Expanding the biomolecular modeling and design toolbox to handle uncertainty As many of the interactions among engineered modules and their context will be difficult to completely disentangle, we will also likely need to embrace uncertainty in the design phase with new modeling and design frameworks that are robust to lack of information, especially in the connectivity among circuit components. In the same way that Thévenin's theorem guides composition of complex electrical circuits under “non-ideal” conditions (without infinite input impedance or zero output impedance),[66] we need tools that allow us to predictably compose modules despite all the uncertainty inherent in biological systems. With this respect, we may also draw inspiration from the course of natural evolution, wherein core processes are repeated and conserved in different contexts, but their connectivity remains fluid and context-dependent through “weak regulatory linkages.”[67] This, too, will require novel approaches; nature has successfully “designed” complex and robust systems via billions of years of trial and error, whereas time is a precious and limited resource in research and technology. In cases where quantitative prediction of system behavior is required in the design process, we may be able to adapt machine learning (ML) approaches to reduce uncertainty in biophysical models and to facilitate system composition and prediction in a variety of contexts. These models could allow us to identify designs that are more likely to perform as intended, thereby reducing the number of circuit variants to be tested and ultimately accelerating the design process. However, our ability to generate strong ML models is tied to data quality: we often have imprecise, indirect, and sparse measurements, wherein only a subset of the system's state can be measured via a proxy, and at times with only population-level resolution. Oftentimes,

the types of measurements we can make depend on the experimental setup, so we need to develop ML models that accommodate these inputs. Furthermore, the outputs of these models also need to be tailored for synthetic biology: ML models, such as neural networks, can have low interpretability, which curtails their utility for informing genetic circuit design. To overcome this, physics-informed ML, developed for other engineering problems, may serve as a starting point.[68]

Conclusions and outlook

Although conceptual analogies between synthetic biology and electrical and computer engineering can play a constructive role in engineering biology, a strict mapping between the fields is not conducive for overcoming current challenges in synthetic biology. In particular, the convenience of applying principles of modular and hierarchical design to engineering biology is undercut by the accompanying insufficient attention paid to the poor robustness to context of today's defined genetic components. If this trend continues, we may not reach a future where engineered cells are employed for myriad real-world applications, for which safety, accuracy, and reliability are paramount. Fundamental research is critically needed to explore novel design approaches that acknowledge context-dependence and achieve emergent robustness from possibly non-robust components. Accordingly, new mathematical formalisms for predicting, exploiting, or mitigating systems' connectivities will be instrumental. Addressing these gaps in our ability to engineer biology will yield a powerful set of tools that will allow us to move synthetic biology and biotechnology forward.

Acknowledgements

Some of the ideas discussed in this paper were generated as part of an NSF-funded workshop under NSF Award #1941841. We want to thank all the workshop participants: Neda Bagheri (U. Washington, Seattle); Enoch Yeung (UCSB); Eduardo Sontag (NEU); Hana El Samad (UCSF); Murat Arcak (UC Berkeley); Brian Munsky (Colorado State U.); Jeff Hasty (UCSD); Howard Salis (Penn State U.); Leonidas Bleris (UT Dallas); Ophelia Venturelli (U. of Wisconsin, Madison); Chris Myers (U. Colorado, Boulder); Marcella Gomez (UC Santa Cruz); Andras Gyorgy (NYU, Abu Dhabi); Diego Oyarzun (U. of Edinburgh (UK)); Xiaojun Tian (Arizona State U.); Xiao Wang (Arizona State U.); Josh Leonard (Northwestern U.); Elisa Franco (UCLA); Mustafa Khammash (ETH); Aura Gimm (OSD); Mitra Basu (NSF); Sankar Basu (NSF); David Ross (NIST); Ron Weiss (MIT); Jordan Berg (NSF).

References

1. Tirard S. La biologie synthétique : de Stéphane Leduc à Craig Vanter… et retour ? Cah Fr Viète 2012;137–150.

2. Meyers M. Assembling, Governing, and Debating an Emerging Science: The Rise of Synthetic Biology in France. BioScience 2013;63:373–379.

3. Monod J, Jacob F. General Conclusions: Teleonomic Mechanisms in Cellular Metabolism, Growth, and Differentiation. Cold Spring Harb Symp Quant Biol 1961;389–401.

4. Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol 1961;3:318–356.

5. Jacob F, Monod J. On the Regulation of Gene Activity. Cold Spring Harb Symp Quant Biol 1961;193–211.

6. English MA, Gayet RV, Collins JJ. Designing Biological Circuits: Synthetic Biology Within the Operon Model and Beyond. Annu Rev Biochem 2021;90:221–244.

7. Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in Escherichia coli. Nature 2000;403:339–342.

8. Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. Nature 2000;403:335–338.

9. Kauffman S. The large scale structure and dynamics of gene control circuits. J Theor Biol 1974;44:167–190.

10. McAdams HH, Arkin A. Simulation of Prokaryotic Genetic Circuits. Annu Rev Biophys Biomol Struct 1998;27:199–224.

11. Endy D. Foundations for engineering biology. Nature 2005;438:449–453.

12. Baker D, Church G, Collins J, et al. Engineering Life: Building a FAB for Biology. Sci Am 2006;294:44–51.

13. Purnick PEM, Weiss R. The second wave of synthetic biology: from modules to systems. Nat Rev Mol Cell Biol 2009;10:410–422.

14. Andries O, Kitada T, Bodner K, et al. Synthetic biology devices and circuits for RNA-based 'smart vaccines': a propositional review. Expert Rev Vaccines 2015;14:313–331.

15. Nielsen AAK, Der BS, Shin J, et al. Genetic circuit design automation. Science 2016;352:aac7341–aac7341.

16. Kitada T, DiAndreth B, Teague B, et al. Programming gene and engineered-cell therapies with synthetic biology. Science 2018;359:eaad1067.

17. Donahue PS, Draut JW, Muldoon JJ, et al. The COMET toolkit for composing customizable genetic programs in mammalian cells. Nat Commun 2020;11:779.

18. Yeung E, Dy AJ, Martin KB, et al. Biophysical Constraints Arising from Compositional Context in Synthetic Gene Networks. Cell Syst 2017;5:11-24.e12.

19. Laboulaye MA, Duan X, Qiao M, et al. Mapping Transgene Insertion Sites Reveals Complex Interactions Between Mouse Transgenes and Neighboring Endogenous Genes. Front Mol Neurosci 2018;11:385.

20. Sauer C, Syvertsson S, Bohorquez LC, et al. Effect of Genome Position on Heterologous Gene Expression in *Bacillus subtilis* : An Unbiased Analysis. ACS Synth Biol 2016;5:942–947.

21. Chen X, Zhang J. The Genomic Landscape of Position Effects on Protein Expression Level and Noise in Yeast. Cell Syst 2016;2:347–354.

22. Cooke K, Browning DF, Lee DJ, et al. Position effects on promoter activity in *Escherichia coli* and their consequences for antibiotic-resistance determinants. Biochem Soc Trans 2019;47:839–845.

23. Kosuri S, Goodman DB, Cambray G, et al. Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*. Proc Natl Acad Sci 2013;110:14024–14029.

24. Normanno D, Boudarène L, Dugast-Darzacq C, et al. Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. Nat Commun 2015;6:7357.

25. Lin L, Liu Y, Xu F, et al. Genome-wide determination of on-target and off-target characteristics for RNA-guided DNA methylation by dCas9 methyltransferases. GigaScience;7 . Epub ahead of print March 1, 2018. DOI: 10.1093/gigascience/giy011.

26. Del Vecchio D, Ninfa AJ, Sontag ED. Modular cell biology: retroactivity and insulation. Mol Syst Biol 2008;4:161.

27. Jayanthi S, Nilgiriwala KS, Del Vecchio D. Retroactivity Controls the Temporal Dynamics of Gene Transcription. ACS Synth Biol 2013;2:431–441.

28. Zhang R, Goetz H, Melendez-Alvarez J, et al. Winner-takes-all resource competition redirects cascading cell fate transitions. Nat Commun 2021;12:853.

29. Frei T, Cella F, Tedeschi F, et al. Characterization and mitigation of gene expression burden in mammalian cells. Nat Commun 2020;11:4641.

30. Jones RD, Qian Y, Siciliano V, et al. An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells. Nat Commun 2020;11:5690.

31. Qian Y, Huang H-H, Jiménez JI, et al. Resource Competition Shapes the Response of Genetic Circuits. ACS Synth Biol 2017;6:1263–1272.

32. Ceroni F, Algar R, Stan G-B, et al. Quantifying cellular capacity identifies gene expression designs with reduced burden. Nat Methods 2015;12:415–418.

33. Gyorgy A, Jiménez JI, Yazbek J, et al. Isocost Lines Describe the Cellular Economy of Genetic Circuits. Biophys J 2015;109:639–646.

34. Nevozhay D, Adams RM, Van Itallie E, et al. Mapping the Environmental Fitness Landscape of a Synthetic Gene Circuit. PLoS Comput Biol 2012;8:e1002480.

35. Shakiba N, Fahmy A, Jayakumaran G, et al. Cell competition during reprogramming gives rise to dominant clones. Science 2019;364:eaan0925.

36. Maheden K, Bashth OS, Shakiba N. Evening the playing field: microenvironmental control over stem cell competition during fate programming. Curr Opin Genet Dev 2021;70:66–75.

37. Liao MJ, Din MO, Tsimring L, et al. Rock-paper-scissors: Engineered population dynamics increase genetic stability. Science 2019;365:1045–1049.

38. Zhang R, Li J, Melendez-Alvarez J, et al. Topology-dependent interference of synthetic gene circuit function by growth feedback. Nat Chem Biol 2020;16:695–701.

39. Nikolados E-M, Weiße AY, Ceroni F, et al. Growth Defects and Loss-of-Function in Synthetic Gene Circuits. ACS Synth Biol 2019;8:1231–1240.

40. McAdams HH, Arkin A. Stochastic mechanisms in gene expression. Proc Natl Acad Sci 1997;94:814–819.

41. Swain PS, Elowitz MB, Siggia ED. Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc Natl Acad Sci 2002;99:12795–12800.

42. Scott M, Gunderson CW, Mateescu EM, et al. Interdependence of Cell Growth and Gene Expression: Origins and Consequences. Science 2010;330:1099–1102.

43. Balázsi G, van Oudenaarden A, Collins JJ. Cellular Decision Making and Biological Noise: From Microbes to Mammals. Cell 2011;144:910–925.

44. Bradley RW, Buck M, Wang B. Recognizing and engineering digital-like logic gates and switches in gene regulatory networks. Curr Opin Microbiol 2016;33:74–82.

45. Fontanarrosa P, Doosthosseini H, Borujeni AE, et al. Genetic Circuit Dynamics: Hazard and Glitch Analysis. ACS Synth Biol 2020;9:2324–2338.

46. Brooks SM, Alper HS. Applications, challenges, and needs for employing synthetic biology beyond the lab. Nat Commun 2021;12:1390.

47. Jessop-Fabre MM, Sonnenschein N. Improving Reproducibility in Synthetic Biology. Front Bioeng Biotechnol 2019;7:18.

48. Tiwari K, Kananathan S, Roberts MG, et al. Reproducibility in systems biology modelling. Mol Syst Biol;17 . Epub ahead of print February 2021. DOI: 10.15252/msb.20209982.

49. Carbonell P, Radivojevic T, García Martín H. Opportunities at the Intersection of Synthetic Biology, Machine Learning, and Automation. ACS Synth Biol 2019;8:1474–1477.

50. Fleur TL, Hossain A, Salis HM. Automated Model-Predictive Design of Synthetic Promoters to Control Transcriptional Profiles in Bacteria. Preprint. Synthetic Biology . Epub ahead of print September 1, 2021. DOI: 10.1101/2021.09.01.458561.

51. Grunberg TW, Del Vecchio D. Modular Analysis and Design of Biological Circuits. Curr Opin Biotechnol 2020;63:41–47.

52. Scott SR, Din MO, Bittihn P, et al. A stabilized microbial ecosystem of self-limiting bacteria using synthetic quorum-regulated lysis. Nat Microbiol 2017;2:17083.

53. Ceroni F, Boo A, Furini S, et al. Burden-driven feedback control of gene expression. Nat Methods 2018;15:387–393.

54. Rackham O, Chin JW. A network of orthogonal ribosome·mRNA pairs. Nat Chem Biol 2005;1:159–166.

55. Bleris L, Xie Z, Glass D, et al. Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template. Mol Syst Biol 2011;7:519.

56. Mishra D, Rivera PM, Lin A, et al. A load driver device for engineering modularity in biological networks. Nat Biotechnol 2014;32:1268–1275.

57. Shopera T, He L, Oyetunde T, et al. Decoupling Resource-Coupled Gene Expression in Living Cells. ACS Synth Biol 2017;6:1596–1604.

58. Huang H-H, Qian Y, Del Vecchio D. A quasi-integral controller for adaptation of genetic modules to variable ribosome demand. Nat Commun 2018;9:5415.

59. Darlington APS, Kim J, Jiménez JI, et al. Dynamic allocation of orthogonal ribosomes facilitates uncoupling of co-expressed genes. Nat Commun 2018;9:695.

60. Lillacci G, Benenson Y, Khammash M. Synthetic control systems for high performance gene expression in mammalian cells. Nucleic Acids Res 2018;46:9855–9863.

61. Segall-Shapiro TH, Sontag ED, Voigt CA. Engineered promoters enable constant gene expression at any copy number in bacteria. Nat Biotechnol 2018;36:352–358.

62. Liu CC, Jewett MC, Chin JW, et al. Toward an orthogonal central dogma. Nat Chem Biol 2018;14:103–106.

63. Aoki SK, Lillacci G, Gupta A, et al. A universal biomolecular integral feedback controller for robust perfect adaptation. Nature 2019;570:533–537.

64. Frei T, Chang C-H, Filo M, et al. Genetically Engineered Proportional-Integral Feedback Controllers for Robust Perfect Adaptation in Mammalian Cells. Preprint. Synthetic Biology . Epub ahead of print December 6, 2020. DOI: 10.1101/2020.12.06.412304.

65. Zhou K, Doyle J, Glover K. Robust and Optimal Control. Prentice Hall. 1996.

66. Thévenin. LC. Extension de la loi d'Ohm aux circuits électromoteurs complexes [Extension of Ohm's law to complex electromotive circuits]. Ann Télégraphiques 1883;10:222–224.

67. Gerhart JC, Kirschner MW. The Plausibility of Life: Resolving Darwin's Dilemma. Yale University Press. 2005.

68. Karniadakis GE, Kevrekidis IG, Lu L, et al. Physics-informed machine learning. Nat Rev Phys 2021;3:422–440.