Build synthetic circuits at different scales

Thesis by
Rongrong Du

In Partial Fulfillment of the Requirements for the
Degree of
Ph.D. in Bioengineering

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2026
Defended Jan 12, 2026

© 2026

Rongrong Du
ORCID: 0009-0003-4942-3020

All rights reserved

ACKNOWLEDGEMENTS

I still remember the late summer of 2021, when I went abroad for the first time. Departing from my wet and rainy spiritual hometown, Shanghai, with excitement, boldness and lingering nostalgia, I was embraced by the irresistible, dazzling Southern California sunshine. With the help of an electronic campus map, I found my way to Ramo Auditorium and arrived twenty minutes early, only to find the entrance seemingly 'locked'. Tom Röschinger, the teaching assistant, opened the door for me and taught me how to properly operate it. Rob Phillips was deriving whether humans or cows on this planet collectively weighed more with a giant chalk in his hand. Several days later, I started to pull my first glass pipette in the Zernicka-Goetz lab.

That was the beginning of my adventure. Looking back at the past four years, I am grateful to Caltech for admitting me as a graduate student and allowing this joyful journey to happen. Here, I have learned to become a self-trained, more mature scientist, both technically and mentally. I have grown more courageous and more honest with myself, no longer hiding my nerdy, quirky hobbies (including diving into geometric problems in my brain in the bathroom), but instead trying to befriend them. Additionally, I learn to value achievements less but cherish training and growth more. I feel immensely fortunate to have glimpsed "an eternal weight of glory beyond all comparison"- I have confirmed with myself that the work I have pursued under the sun is not meant to nourish my own ego, but to plunge into the unknown and, through the spreading of knowledge, ultimately contribute to the common good.

I feel profoundly grateful to have worked with my PhD advisor, Michael Elowitz. As every trainee in our lab has described, he is undoubtedly open-minded, generous, inspiring, and supportive. But on a more personal level, I want to record several conversations with him that I believe will stay with me for the rest of my life.

Earlier this year, when I began thinking about my career path after graduation, I pressured myself to identify one big question that I could devote my entire life to. Unsurprisingly, I failed - there are too many exciting questions and too little time in a human lifespan. Therefore, I switched to a different strategy: treating the decision like a "maximum-optimization" problem. I tried to account for everything: how academia might judge different training paths, how fundable a question might

be, and many other external constraints. The more I tried to optimize, the more frustrated I became.

At that time, I talked with Michael. What he told me is something I will remember for the rest of my life: that I should not try to fit into the frame set by the world, but instead pursue what is new and fun, and let the world adjust itself to welcome Rongrong's ideas, creations, and style. From that moment, I started to believe that I could one day establish my own way of doing science, and perhaps even influence the next generation of scientists.

Yet some time later, after being judged, defeated, and watching people struggle through the faculty-job landscape, I felt discouraged again, especially when I began searching for postdoctoral positions myself. One day I complained to Michael that academia seemed to give young scientists no room for mistakes. He disagreed. He reminded me that mistakes are not only allowed but inevitable—that even his friends in their 50s and 60s make mistakes all the time. He said that the polished lives presented on social media are far from the truth, and that I should simply feel fearless and go wherever I genuinely want to go.

Michael, I want to tell you that it has been one of the great fortunes of my life to inherit this wisdom about being a great scientist from you.

I also want to express my most sincere gratitude to several other faculty members at Caltech. I am deeply thankful to Justin Bois for imparting so much of his knowledge. His courses were always well-designed and wonderfully dense, making me feel as though I were a soldier being trained in the armies of mathematics and physics—and, more importantly, I truly enjoyed the process. Whenever I worked on Justin's problem sets, I found myself wholly absorbed; it felt like wandering through a vast puzzle castle, knowing that at least one path to the solution must exist, and that a smiling magician would await me at the gate once I found it. Because of Justin and his classes, I gradually became more confident, even audacious, in taking up analytical and theoretical challenges.

I am also grateful to Justin for inviting me twice to serve as a teaching assistant for BE150, an opportunity I could only accept once due to travel. Before TA-ing the course, I knew only that I loved teaching and felt a strong desire to 'pass the torch.' Afterward, I discovered not only that I empathized deeply with students (and sometimes even 'parent' them), but also that my ego in teaching had been humbled. I came to realize that teaching and mentoring resemble planting trees: I can plant

and water with care, but each tree grows at its own pace. I admire Justin immensely for understanding the growth rates of individual trees, designing the most suitable ways to nurture them, and yet always remaining fair.

I am truly appreciative of the support offered by many other faculty members. I want to thank Viviana Gradinaru for her trust and the resources she generously provided. I am grateful to Magdalena Zernicka-Goetz, who welcomed me as a rotation student and opened the door of developmental biology to me. I thank Marianne Bronner, my committee chair, for always being available to talk about science and career decisions, and for giving me direct and honest advice whenever I asked. I am thankful for the developmental biology course co-taught by Magda, Marianne, and Ellen. That course revealed to me the profound beauty of developmental biology. To me, the developmental process is like a masterfully orchestrated symphony. The course, together with the experiences of working with great developmental biologists, inspired me to challenge myself once again: to step outside my own conceptual world, embrace a mechanistic mindset, and ultimately join a neurodevelopment lab for my postdoctoral research.

I also want to thank Matt Thomson and Long Cai for serving on my thesis committee, and Mitch Guttman and Alexei Aravin for being members of my first-year advisory committee.

I am grateful to all my wonderful collaborators over the years for their dedication and generosity. First and foremost, I want to thank Michael J. Flynn and Sheng Wang for generously sharing their materials, expertise, and time on the miRNA circuit project and the patterning project, respectively. Mike was the very first person I worked with in the Elowitz lab; he welcomed me as a rotation student and showed me around the lab. I have always admired Mike as a person – I'd say he is a true Gryffindor. His courage, honesty, and unwavering commitment to integrity have encouraged me to remain steady when facing the turbulence of the world.

Sheng, on the other hand, is what I would call a pure scientist. I still remember our long conversation on Halloween night three years ago, when he described the Turing pattern problem as 'gorgeous' and said that even if he could never fully solve it, he would still consider it gorgeous. His words always remind me of Confucius's saying: "Though I cannot reach it, my heart yearns for it."

Beyond our scientific collaborations, I have also become close friends with both Mike and Sheng. We have spent countless hours hanging out, supporting each other

through both work and life.

In addition to Mike and Sheng, I want to thank Karan Mahe for boldly taking up the challenge of delivering and monitoring the circuit in the live mice brain. Rana Eser, and the gene therapy subgroup of the Gradinaru lab continuously endeavored on utilizing the circuit to develop tightly controlled gene therapy in the future. JS Paul co-TA-ed with me for BE150, and it was a great pleasure to have worked with him. Inna Strazhnik has been always super-efficient and made fantastic illustrations for all my papers (and thank you for sharing with me your life wisdom, Inna!). Outside of Caltech, Monique Honsa and her advisor Ralf Jungmann at Max Planck Institute of Biochemistry have generously resourced us with DNA-PAINT technology and tested our circuit in imaging the membrane receptors. Sean E. McGeary at Harvard Medical School has provided valuable and constructive feedback on the miRNA biology experiments. Cindy Ren at Eastern Institute of Technology, Ningbo, has organized multiple exciting discussions on the circuit and noise.

I almost cried when I began writing this part, which is meant to honor the mentors and friends who shaped my time at Caltech. My first mentor was Melissa E. Truong. My best moment in the Zernicka-Goetz lab was spent sitting with Melissa in the microscopy room, watching immunostained preimplantation embryos glow against the dark screen. Before rotating in the lab, I had never worked with animals, and people often said that mouse embryos were among the most challenging to handle. To prepare me for this work, Melissa thoughtfully designed my entire training process. She first familiarized me with the anatomy of the mouse reproductive system, then guided me through dissections using sacrificed mice, and finally asked me to shadow her to the animal facility. I went from being nervous about animal work to feeling calm, experienced, and full of respect while performing it - and none of that growth would have been possible without Melissa. Beyond the technical training, she deeply influenced me in what it means to be a rigorous biologist. Moreover, I must also say that when I started my rotation, I was still in the middle of a culture shock, since it was the first lab that I rotated with. Melissa was extremely considerate, almost like an older sister in the lab - helping me communicate and find a sense of belonging.

During my rotation in the Zernicka-Goetz lab, Elena R. Kingston joined us, and I was incredibly fortunate to gain another kind and capable “big sister.” Elena helped me set up the embryonic stem cell culture system, and after I moved on to the miRNA project in the Elowitz lab, I continued to seek her advice, since she was

also an expert in miRNA biology. I also wanted to thank Meng, Min, Dong-Yuan, Kate, Clare, Victoria, and Nia for their advice, encouragement, and companionship during my time in the Zernicka-Goetz lab.

It has truly been a gift to meet Dongyang Li in the Elowitz lab. Although we never deeply collaborated on my major projects, he has unquestionably been a real mentor to me. Dongyang was my office neighbor, which made it easy for me to ask questions, and luckier still, we happened to share a similar mindset, skill set and even set of values. The difference was that he had already taken the risks and accumulated wisdom, while I was fortunate enough to learn from his experiences. I have countless cherished memories with him; if I were to write them all down, they would easily fill several pages. But there is one recent moment I especially want to record, in case one day I or someone reading this part needs a reminder.

Back in September, I received all kinds of advice about finding a postdoctoral position. I knew each suggestion came from a place of wanting me to succeed. Yet I felt immense pressure, because what I truly wanted was not 'success,' but the pure, endless pursuit of the beauty of science itself. Still, it seemed that without achieving a conventional form of success, I might have little chance to pursue my favorite questions in such a competitive world. I felt torn: part of me wanted to retreat into my own utopia rather than compromise; another part wondered if I might still have the strength to fight another round. I texted my brokenness to Dongyang, and he sent me an image from *The Boy and the Heron* and told me that if externally validated success was not what I cared about, then I should try to be brave and ignore those constraints. I still feel proud that I followed that advice.

Even more fortunately, we became close friends outside the lab. People often joked that Dongyang and his wife, Shihan, were my "parents" in the lab. They took me to our favorite Joe Hisaishi concert and once gave me a Mandrake LEGO set as a holiday gift. During the fierce mountain fire this year, they promised they would take me with them no matter where I was - and I trusted them completely. It was deeply comforting to have them as my family here.

I still cannot believe how lucky I was to gain yet another pair of "lab parents": James Linton and Leah Santat. Everyone in our lab must have thanked James and Leah in their thesis acknowledgements for taking care of the lab and supporting us administratively - and I could not agree more. But I want to say something additional.

I like to think of James as a true *shi* in the sense of ancient Chinese culture—a person who willingly walks the "thorny road of honor." James devoted a tremendous amount of effort to helping us bring the patterning project to completion. But even more importantly, he teaches us through the weight of his own life. After becoming friends with James, I felt convinced that lifelong learning is not only possible but deeply admirable. Once, I gave him a science-fiction book, *The Wandering Earth*, by a Chinese author. James read it very carefully and even took notes - just so he could discuss it with us. One of the sweetest compliments I received during my years at Caltech came from James: he called me an owl, his favorite bird, symbolizing wisdom and knowledge in western culture. I still treasure that.

Leah, meanwhile, was another one of my office neighbors. Naturally, we gossiped every day, shared snacks, and took central control of the lab's food supply. Whenever Leah returned from a holiday trip, I would find a small gift sitting on my desk—and I made sure she would find one, too.

Please forgive me for being unable to fully articulate the depth of my gratitude to all the postdocs and senior graduate students who have offered me countless pieces of invaluable advice over the years. Yujie Fan showed me what it means to be an independent, outstanding, and truly beautiful female scientist. Fangzhou Xiao, my assigned bioengineering big sibling, taught me how to think like a knowledgeable and reflective theorist. He is also the kind of exceptional teacher I hope to become one day—passionate, patient, and deeply committed to sharing everything he knows with his students, who can undoubtedly feel that dedication. Bo Gu is a master of sarcasm, and through his humor I learned to be more honest with myself. Shiyu Xia, Yodai Takei, Roni Hadas, Zibo Chen, Judy Shon, Maire Gavagan, Felix Horns, Ronghui Zhu, Yitong Ma, Lucy Chong, Huan Edmonds, Duncan Chadly, Martin Tran—and of course our devoted administrators, Rui Malinowski and Jo Leonardo—thank you all for contributing to the long, intricate process of shaping me into a more mature scientist.

In addition to having the privilege of standing on the shoulders of giants, I want to thank my peers for making the journey joyful on the ground. I feel incredibly lucky to have met my office mate, Kaiwen Luo, who entered graduate school the same year as I did. People often say that a true Hufflepuff is rare in real life, especially in a world where individualism prevails and self-realization and personal achievement often overshadow integrity, dedication, and care for others. But Kaiwen is, without question, a real Hufflepuff. In science, he combines remarkable engineering talent

with the quiet tenacity of a craftsman. In life, he chooses kindness over judgment every time. These qualities have inspired me deeply. Debating scientific questions with Kaiwen has always been a great joy—I never worried about offending him, because I knew he cared only about the truth, and that sense of security is rare and precious. And in everyday life, we could always rely on Kaiwen, simply because he is so unfailingly reliable.

I also want to thank all my other friends, both inside and outside the lab. My quality of life improved dramatically when Xuanjia Ye moved from Princeton to Caltech. Chuqi Lu and Hong Han cooked for me and escorted me to work when I broke my knee. Evan Mun, Victoria Tobin, Jan Gregrowicz, Joe Martinez, Mehernaz Haque, Leslie Klock, Ali Ahmed Diaz, Haojing Chen, Hengyu Li, Haowen Zhou, Ziyan Wu, Qinyu Han, Lingjing Fang, Can Li, Xun Wang, Changfan Lin, Manxuan Zhou, Yunqing Wang, Canran Wang—and many other friends who I owed but forgot—thank you for filling my years here with so many wonderful memories.

Last but not least, I want to express my deepest gratitude to my family and to my closest friends from before Caltech, who have been the bedrock of my life. I thank my parents for giving me both nature and nurture. Only as I grew older did I truly realize how extraordinary they are. I am incredibly fortunate to have parents who are also my genuine friends—we video call almost every day, sharing what we have seen, experienced, and thought about. I cherish the fact that they regard me as an independent individual rather than an extension of themselves; they respect every choice I make while supporting me with everything they can offer.

Special thanks go to my uncle's family in Canada. The first package I received after arriving in the U.S. was a giant Instant Pot from him—though I confess I did not cook with it nearly as often as I should have. My cousin and cousin-in-law have always been generous and loving, sometimes driving for hours just to visit me, supporting me both materially and emotionally. My extended family in Shanghai and Shanxi always welcomed me home with warmth every time I returned.

It was a beautiful stroke of serendipity that I went to Fudan University for my undergraduate study and met a group of people who I know will be lifelong friends. Jingyi Hu and I have known each other since freshman year. We shared the same major, joined the same lab, worked, ate, and wandered through life side by side, witnessing each other's growth. Although we attended different graduate schools, we never lost touch—through daily messages and annual reunions. Jingyi is a giver who takes care of everyone around her. I want to thank her for tolerating my quirky

metaphors, blunt criticisms, and occasionally harsh judgments—and for keeping the good parts. I felt secure enough with her to be both loyal and honest. Beyond Jingyi, I also want to thank my other college friends—Shengyi Lu, Fangfei Ye, Ziyan Yu, Yuyan Zeng, Chuyue Zhang, Zixuan Zhang, Jianghui Zhu, Yang Xie, Chen Li, Qinyi Ma, Yuzhe Gu, Yunzhi Lin, Jie Wang, Jiahe Chen, Jiexin Shen, Xuexia Duan, Wenxi Li, Meng Ye, Zijin Chen, and many others whose names I may have forgotten but whose kindness remains vivid. Every time we talked, met, or even crossed my mind, I felt warm and strengthened.

Another precious gift in my life is meeting Sophie. As I came to know her, I learned what it means for inner and outer beauty to exist in seamless unity. Sophie has been—and will continue to be—my mentor, confidant, and kindred spirit. Every day we talk about everything from funny stories about my lab mates to anecdotes about historical figures. Our shared curiosity about each other's academic worlds gives us the steepest, happiest learning curves. Many of the moments we've shared are too sentimental to record here, but I want to preserve one conversation from this summer. As I mentioned above, during a period when I felt dispirited by current academia, I confided in Sophie. She told me that *we can let your success define what success is.* It was the greatest trust anyone has ever placed in me, and it encouraged me immeasurably. We have been, and will continue to be, bound by shared ideals and a common vision. I could never have imagined how good my life would become from the moment I met you, Sophie.

Finally, I would like to close this acknowledgement with a line of thought inspired by my recent favorite song, *Advice From a Caterpillar*. As I take the next step, I will carry with me the courage to take risks and the willingness to wander into places no one has yet explored.

ABSTRACT

Multicellular organisms rely on the coordinated actions of diverse organs to sustain life. Each organ comprises cells that communicate with each other to execute physiological functions, and each cell encodes gene regulatory networks that shape its gene expression programs. The intrinsic complexity of biological systems, including features such as redundancy that endow them with robustness, also makes them difficult to study using reductionist approaches alone.

To elucidate quantitative design principles underlying multicellular organization, I adopted a bottom-up approach and built synthetic circuits at multiple scales. In the first project, I engineered a single-gene incoherent feedforward circuit that leverages multispecific microRNA targeting to achieve dosage-invariant and tunable protein expression across wide ranges of gene copy numbers. In the second project, I constructed a multicellular reaction–diffusion circuit that integrates juxtacrine and paracrine signaling to generate self-organized, periodic Turing patterns.

Together, these studies introduce new tools for engineering regulatory behaviors, reveal general principles that govern biological organization across scales, and pave the way for potential translational applications.

PUBLISHED CONTENT AND CONTRIBUTIONS

[1] Rongrong Du, Michael J. Flynn, Karan Mahe, Monique Honsa, Bo Gu, Dongyang Li, Sean E. McGeary, Viviana Gradinaru, Ralf Jungmann, and Michael B. Elowitz. “miRNA modules for precise, tunable control of gene expression”. In: *Molecular Cell* 86(1) (2026), 194–212.e7. DOI: 10.1016/j.molcel.2025.11.028.

R.D., M.F., and M.B.E. conceived and designed the study. M.B.E. supervised the study. R.D., M.F., K.M., M.H., B.G., D.L., and S.E.M. performed or assisted with experiments and data analysis. R.D. and M.B.E. wrote the manuscript with input from all authors.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

INTRODUCTION

1.1 The beauty of the bottom-up approach

As early as 1961, Monod and Jacob pioneered the concept of the *gene circuit* through their operon model, and further envisioned whether regulatory mechanisms identified in bacteria could be extended to the tissues of higher organisms [1]. Rather than focusing on individual molecular components contributing to *E. coli* physiology, they proposed that "their known elements could be connected into a wide variety of 'circuits', endowed with any desired degree of stability" [2]. This vision captures the central idea of the bottom-up approach: by assembling minimal, well-defined elements into synthetic circuits, one can directly test how abstract regulatory motifs give rise to molecular, cellular, and even multicellular behaviors.

Over the past decades, synthetic biology has increasingly embraced this bottom-up strategy to uncover quantitative design principles of biological systems. In the early 2000s, the construction of the first synthetic oscillator and bistable toggle switch in bacteria demonstrated that it is regulatory circuitry—rather than the specific molecular identity of its components—that underlies desired dynamical functions [3, 4]. Since then, the rapid development of molecular and cellular engineering tools has enabled bottom-up approaches across multiple biological scales. At the molecular level, transcription has been revealed as a dynamic process whose frequency, amplitude, and duration can be precisely tuned to encode functional outcomes [5, 6]. At the cellular and multicellular levels, signaling pathways have been reframed as a shared "language" that supports promiscuous and combinatorial interactions [7, 8, 9]. At the organismal level, it has become clear that circuit architecture, rather than molecular identity, can dominate patterning outcomes—for example, GFP can functionally replace the morphogen Dpp in fly wing patterning when embedded in the appropriate regulatory circuitry [10].

Compared to traditional top-down approaches, the bottom-up approach offers several distinct advantages.

First, it enables more direct and quantitative mappings between inputs and outputs. Natural gene regulatory networks are intrinsically intertwined and highly redundant [11], such that perturbation of one or even multiple genes can be masked

by endogenous compensatory mechanisms. In contrast, bottom-up designs often operate in relatively isolated contexts using components that are orthogonal to endogenous networks, allowing clean measurements and predictable quantitative behaviors defined by circuit topology.

Second, the bottom-up approach circumvents evolutionary constraints by leveraging rationally engineered components. As François Jacob famously noted, natural selection acts as a tinkerer, assembling workable solutions rather than designing optimally from first principles [12]. Bottom-up strategies complement this process by enabling modular and systematic interrogation of evolution-selected traits. For instance, advances in large-scale genome assembly have made it possible to study how the colinearity of mammalian Hox gene clusters is regulated during development through controlled perturbations of circuit organization [13].

Third, bottom-up approaches allow the exploration of novel functions through rewiring of naturally evolved pathways. By converting an endogenous toggle switch into a sustained oscillator, yeast cells exhibit extended lifespan due to delayed commitment to aging, illustrating how new dynamical regimes can lead to emergent phenotypes [14].

Finally, bottom-up designs are inherently extensible, enabling both scalability toward increased functional complexity and portability across biological contexts. For example, synthetic multistable systems can be expanded to additional stable states simply by introducing new transcription factors, without re-engineering the existing circuit architecture [15]. Such designs provide a conceptual framework for generating diverse cell types or lineages relevant to developmental processes. Similarly, the promiscuous signaling framework—originally developed to explain combinatorial BMP signaling—has since been generalized to other systems, including GPCR signaling [16] and multicomponent surface condensates [17].

Taken together, the bottom-up approach provides a uniquely powerful and elegant framework for probing fundamental biological principles by transforming abstract regulatory hypotheses into experimentally testable systems.

1.2 Single-gene regulatory circuits

Despite the overall robustness of gene expression landscapes, the expression level of a single gene can be functionally critical in specific biological contexts. Human genetics provides abundant evidence for sensitivity to transcription factor dosage [18]. For instance, modest changes in SOX9 protein levels within facial

progenitor cells are sufficient to alter chondrogenesis and ultimately drive craniofacial shape variation [19]. Similarly, in multiple neurological disorders—including Smith–Magenis syndrome [20], Rett syndrome [21], and MECP2 duplication syndrome [22, 23]—haploinsufficiency or triplosensitivity of a single gene is sufficient to cause severe and often lethal phenotypes. Together, these observations highlight the importance of quantitatively understanding regulatory mechanisms that govern single-gene expression.

From an engineering perspective, single-gene systems also offer a uniquely tractable platform for bottom-up investigation. A single gene reporter provides a clear, quantitative readout that enables precise mapping between regulatory inputs and expression outputs. Over the past decade, a diverse set of engineered molecular tools has been developed using single-gene reporter frameworks to interrogate gene regulation at multiple layers, from transcriptional control to chromatin state. For example, programmable zinc finger–based transcriptional regulators enabled systematic exploration of transcriptional logic in a minimal setting [24]. Building on similar single-gene platforms, subsequent studies revealed how transcription factors combinatorially regulate gene expression through cooperative assembly and higher-order interactions [25]. Likewise, single-gene reporters have been instrumental in quantitatively characterizing the dynamics and functional consequences of epigenetic regulators [26, 27].

Taken together, single-gene regulatory circuits represent a minimal yet powerful abstraction for studying how quantitative control of gene expression gives rise to biological function. Continued expansion of single-gene toolkits is therefore poised to enable broad biological discovery while providing a rigorous foundation for the design of more complex regulatory systems.

1.3 Multicellular circuits and synthetic developmental biology

Cells communicate with one another to transmit information and coordinate complex physiological functions. As a result, multicellular circuits are ubiquitous across developmental stages and evolutionary lineages. Reconstructing multicellular circuits in a synthetic context offers a powerful route to explore alternative developmental outcomes, thereby deepening our understanding of developmental processes and providing new conceptual and practical foundations for disease modeling, regenerative medicine, and biopharmaceutical applications [28, 29].

A defining feature of living systems is their capacity for self-organization: starting

from a single cell, organisms reliably assemble into structured, functional bodies through genetically encoded rules. Synthetic multicellular circuits seek to encode a minimal set of instructions sufficient for collective systems to recapitulate desired developmental architectures. Central questions naturally arise from this endeavor. What regulatory mechanisms govern the emergence of specific spatial organizations, and is there a minimal circuitry capable of generating a given architecture? How robust and flexible are these architectures—do they adapt to external perturbations by transitioning into alternative structures, or do they resist change to preserve form? How do multicellular architectures evolve in space and time, and to what extent can bottom-up method-established structure in vitro recapitulate the dynamics of in vivo development?

Over the past decade, synthetic developmental biologists have developed diverse bottom-up strategies to interrogate multicellular circuits. One approach involves reusing, rewiring, or reconstituting endogenous morphogen systems in simplified synthetic settings. For example, Matsuda et al. demonstrated how lateral inhibition through the Notch–Delta system drives cell differentiation in CHO cells lacking endogenous Notch and Delta expressions [30]. Sekine et al. subsequently constructed Nodal–Lefty–regulated Turing patterns in HEK293 cells devoid of endogenous signaling components [31]. Similarly, Li et al. reconstituted Sonic Hedgehog signaling gradients in NIH 3T3 cells, which transduce Hedgehog signals without undergoing differentiation and do not naturally express Hedgehog ligands [32].

A complementary strategy engineers fully synthetic components to wire multicellular regulatory motifs. Notably, Toda et al. constructed layered and self-sorted multicellular structures using synthetic Notch receptor–controlled interactions [33]. More recently, the development of programmable synthetic adhesion molecules has enabled systematic investigation of cell–cell interactions and spatial organization in self-organizing assemblies [34, 35].

Beyond purely molecular circuits, multicellular self-organization can also be shaped by integrating mechanical constraints [36, 37], chemical and molecular signaling cues [38, 39], and known gene regulatory networks [40]. Rather than constructing fully synthetic systems, these approaches leverage engineered environments—such as geometric confinement or controlled signaling landscapes—to create *ex vivo* architectures that reveal the operating principles of developmental self-organization.

Looking forward, the integration of synthetic genetic circuits, engineered signaling pathways, and physical constraints is expected to enable increasingly faithful re-

constructions of developmental processes. Such hybrid strategies promise not only deeper mechanistic insight into self-organization, but also next-generation tissue engineering approaches that more closely align with physiological regimes.

1.4 Content of this thesis

This thesis presents two synthetic regulatory circuits that operate at distinct biological length scales, illustrating how bottom-up design principles can be applied from single genes to multicellular systems. In Chapter 2, I describe a single-gene regulatory circuit that achieves dosage compensation, demonstrating how precise quantitative control of gene expression can be encoded within a minimal regulatory architecture. In Chapter 3, I present a multicellular circuit that generates reaction–diffusion patterns, showing how local regulatory interactions can give rise to emergent spatial organization at the tissue scale.

Together, these two systems exemplify a unified bottom-up framework for probing regulatory principles across scales, from molecular-level control to multicellular self-organization.

References

[1] Jacques Monod and François Jacob. *General Conclusions: Teleonomic Mechanisms in Cellular Metabolism, Growth, and Differentiation*. 1961.

[2] F Jacob and J Monod. “On the regulation of gene activity”. In: *Cold Spring Harb. Symp. Quant. Biol.* 26.0 (Jan. 1961). Publisher: Cold Spring Harbor Laboratory, pp. 193–211.

[3] M B Elowitz and S Leibler. “A synthetic oscillatory network of transcriptional regulators”. In: *Nature* 403.6767 (Jan. 2000), pp. 335–338.

[4] T S Gardner, C R Cantor, and J J Collins. “Construction of a genetic toggle switch in Escherichia coli”. In: *Nature* 403.6767 (Jan. 2000), pp. 339–342.

[5] Yihan Lin et al. “Combinatorial gene regulation by modulation of relative pulse timing”. In: *Nature* 527.7576 (Nov. 2015), pp. 54–58.

[6] Ravi V Desai et al. “A DNA repair pathway can regulate transcriptional noise to promote cell fate transitions”. In: *Science* 373.6557 (Aug. 2021).

[7] Yaron E Antebi et al. “Combinatorial Signal Perception in the BMP Pathway”. In: *Cell* 170.6 (Sept. 2017), 1184–1196.e24.

[8] Heidi E Klumpe et al. “The context-dependent, combinatorial logic of BMP signaling”. In: *Cell Syst* 13.5 (May 2022), 388–407.e10.

[9] Christina J Su et al. “Ligand-receptor promiscuity enables cellular addressing”. In: *Cell Syst* 13.5 (May 2022), 408–425.e12.

[10] Kristina S Stapornwongkul et al. “Patterning and growth control in vivo by an engineered GFP gradient”. In: *Science* 370.6514 (Oct. 2020), pp. 321–327.

[11] D Hanahan and R A Weinberg. “The hallmarks of cancer”. In: *Cell* 100.1 (Jan. 2000), pp. 57–70.

[12] Francois Jacob. *Evolution and tinkering*. 1977.

[13] Sudarshan Pinglay et al. “Synthetic regulatory reconstitution reveals principles of mammalian cluster regulation”. In: *Science* 377.6601 (July 2022), eabk2820.

[14] Zhen Zhou et al. “Engineering longevity-design of a synthetic gene oscillator to slow cellular aging”. In: *Science* 380.6643 (Apr. 2023), pp. 376–381.

[15] Ronghui Zhu et al. “Synthetic multistability in mammalian cells”. In: *Science* 375.6578 (Jan. 2022), eabg9765.

[16] Andrew B Kleist et al. “Encoding and decoding selectivity and promiscuity in the human chemokine-GPCR interaction network”. In: *Cell* 188.13 (June 2025), 3603–3622.e27.

[17] Aidan Zentner et al. “Information processing driven by multicomponent surface condensates”. In: *arXiv [physics.bio-ph]* (2025).

[18] J G Seidman and Christine Seidman. “Transcription factor haploinsufficiency: when half a loaf is not enough”. In: *J Clin Invest* 109.4 (Feb. 2002), pp. 451–455.

[19] Sahin Naqvi et al. “Precise modulation of transcription factor levels identifies features underlying dosage sensitivity”. In: *Nat Genet* 55.5 (May 2023), pp. 841–851.

[20] Wei-Hsiang Huang et al. “Molecular and Neural Functions of Rai1, the Causal Gene for Smith-Magenis Syndrome”. In: *Neuron* 92.2 (Oct. 2016), pp. 392–406.

[21] B Hagberg et al. “A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett’s syndrome: report of 35 cases”. In: *Ann Neurol* 14.4 (Oct. 1983), pp. 471–479.

[22] Ann L Collins et al. “Mild overexpression of MeCP2 causes a progressive neurological disorder in mice”. In: *Hum. Mol. Genet.* 13.21 (Nov. 2004), pp. 2679–2689.

[23] Melissa B Ramocki et al. “Autism and other neuropsychiatric symptoms are prevalent in individuals with MeCP2 duplication syndrome”. In: *Ann Neurol* 66.6 (Dec. 2009), pp. 771–782.

[24] Ahmad S Khalil et al. “A synthetic biology framework for programming eukaryotic transcription functions”. In: *Cell* 150.3 (Aug. 2012), pp. 647–658.

[25] Caleb J Bashor et al. “Complex signal processing in synthetic gene circuits using cooperative regulatory assemblies”. In: *Science* 364.6440 (May 2019), pp. 593–597.

[26] Lacramioara Bintu et al. “Dynamics of epigenetic regulation at the single-cell level”. In: *Science* 351.6274 (Feb. 2016), pp. 720–724.

[27] Josh Tycko et al. “High-Throughput Discovery and Characterization of Human Transcriptional Effectors”. In: *Cell* 183.7 (Dec. 2020), 2020–2035.e16.

[28] Eric H Davidson, David R McClay, and Leroy Hood. “Regulatory gene networks and the properties of the developmental process”. In: *Proc Natl Acad Sci U S A* 100.4 (Feb. 2003), pp. 1475–1480.

[29] Onur Aydin et al. “Principles for the design of multicellular engineered living systems”. In: *APL Bioeng* 6.1 (Mar. 2022), p. 010903.

[30] Mitsuhiro Matsuda et al. “Synthetic lateral inhibition governs cell-type bifurcation with robust ratios”. In: *Nat Commun* 6 (Feb. 2015), p. 6195.

[31] Ryoji Sekine, Tatsuo Shibata, and Miki Ebisuya. “Synthetic mammalian pattern formation driven by differential diffusivity of Nodal and Lefty”. In: *Nat Commun* 9.1 (Dec. 2018), p. 5456.

[32] Pulin Li et al. “Morphogen gradient reconstitution reveals Hedgehog pathway design principles”. In: *Science* 360.6388 (May 2018), pp. 543–548.

[33] Satoshi Toda et al. “Programming self-organizing multicellular structures with synthetic cell-cell signaling”. In: *Science* 361.6398 (July 2018), pp. 156–162.

[34] George Chao et al. “helixCAM: A platform for programmable cellular assembly in bacteria and human cells”. In: *Cell* 185.19 (Sept. 2022), 3551–3567.e39.

[35] Adam J Stevens et al. “Programming multicellular assembly with synthetic cell adhesion molecules”. In: *Nature* 614.7946 (Feb. 2023), pp. 144–152.

[36] Tony Y-C Tsai et al. “An adhesion code ensures robust pattern formation during tissue morphogenesis”. In: *Science* 370.6512 (Oct. 2020), pp. 113–116.

[37] Sichen Yang et al. “Morphogens enable interacting supracellular phases that generate organ architecture”. In: *Science* 382.6673 (Nov. 2023), eadg5579.

[38] Marina Sanaki-Matsumiya et al. “Periodic formation of epithelial somites from human pluripotent stem cells”. In: *Nat Commun* 13.1 (Apr. 2022), p. 2325.

[39] Yuchuan Miao et al. “Reconstruction and deconstruction of human somitogenesis in vitro”. In: *Nature* 614.7948 (Feb. 2023), pp. 500–508.

[40] Bailey A T Weatherbee et al. “Pluripotent stem cell-derived model of the post-implantation human embryo”. In: *Nature* 622.7983 (Oct. 2023), pp. 584–593.

Chapter 2

SINGLE-GENE CIRCUIT: MIRNA MODULES FOR PRECISE, TUNABLE CONTROL OF GENE EXPRESSION

2.1 Summary

Accurate control of transgene expression is important for research and therapy but challenging to achieve in most settings. miRNA-based regulatory circuits can be incorporated within transgenes for improved control. However, the design principles, performance limits, and applications of these circuits in research and biotechnology have not been systematically determined. Here, combining modeling and experiments, we introduce miRNA-based circuit modules, termed DIMMERs, that establish precise, tunable control of transgene expression across diverse cell types to facilitate imaging, editing, and gene therapy. The circuits use multivalent miRNA regulatory interactions to achieve nearly uniform, tunable, protein expression over two orders of magnitude variation in gene dosage. They function across diverse cell types, and can be multiplexed for independent regulation of multiple genes. DIMMERs reduce off-target CRISPR base editing, improve single-molecule imaging, and allow live tracking of AAV-delivered transgene expression in mouse cortical neurons. DIMMERs thus enable accurate regulation for research and biotechnology applications.

2.2 Introduction

Biomedical research and biotechnology heavily rely on transgene expression in living cells. The ability to accurately establish transgene expression at desired levels is critically needed in many contexts. For example, in gene therapy, overexpression of therapeutic transgenes can be toxic [1]. Similarly, in gene editing and imaging applications, overexpression can reduce specificity or increase background, respectively. However, popular expression systems, including DNA transfection and AAV vectors, as well as integrating systems such as lentivirus [2] or piggyBac transposons [3], typically generate a broad range of expression levels, due to variability in the number of gene copies taken up, integrated, and expressed by each cell, as well as gene expression noise [4, 5]. Selecting individual stable clones can reduce variability but is time-consuming and impossible for gene therapy. What is needed is a simple gene regulation system that could compensate for unavoidable variation

in delivery and expression (**Figure 2.1A**). The ideal system should have several key features: First, it should be genetically compact for compatibility with most delivery vectors. Second, it should allow predictive tuning of the expression setpoint. Third, it should permit construction of multiple independent (orthogonal) regulation systems for simultaneous control of multiple genes. Finally, it should operate across multiple cell types (portability) (**Figure 2.1B**). The incoherent feed-forward loop (IFFL) circuit motif provides an ideal foundation for these capabilities [6]. A circuit in which a target gene and its negative regulator are both encoded in the same DNA construct represents an IFFL-like configuration in which gene dosage, considered as an input, modulates expression of both the target and its negative regulator. When these effects effectively cancel out, target expression can asymptotically approach a constant expression level independent of gene dosage (**Figure 2.1C**).

Previous studies have introduced synthetic circuits based on this principle. Bleris et al. showed that an IFFL based on microRNA (miRNA hereafter) as the negative regulator could achieve dosage compensation [7]. Strovas et al. introduced a similar design incorporating a natural miRNA and multiple repeats of its binding site within the target gene, and examined its expression dynamics over several days [8], achieving dosage compensation over a $\sim$ 20-fold range at the cost of potential crosstalk with endogenous genes. Yang et al later introduced an "equalizer" architecture that combined transcriptional negative feedback through the TetR protein with feed-forward miRNA regulation [9]. This generated an extended regime of strong dosage compensation but required expression of a bacterial protein. Tradeoffs between circuit complexity and efficiency are worth exploring further. Finally, while this work was under review, Love et al demonstrated other configurations of miRNA-based IFFL circuits for dosage compensation [10]. Nevertheless, fundamental questions have remained unclear: What sequence and circuit design principles optimize dosage compensation? Can these systems allow expression tuning, multiplexing of independent regulatory systems in the same cell, and portability across cell types and delivery modalities? And can they provide durable control in vivo for gene therapy applications? To address these goals, we combined mathematical modeling, synthetic design, and quantitative circuit analysis to create a set of miRNA-based dosage compensation systems termed DIMMERs (Dosage Invariant miRNA-Mediated Expression Regulators). These circuits take advantage of multivalent miRNA regulation through the natural TNRC6 scaffold system [11]. They allow systematic tuning of expression levels by modulating the number of miRNA cassettes, numbers of target binding sites, and miRNA-target site complementarity.

Further, they can be used to orthogonally regulate multiple genes in the same cell, and operate similarly across different cell types. A toolkit of ten mutually orthogonal ready-to-use expression systems can be incorporated into diverse systems. They facilitate biological imaging, improve CRISPR base editing, and function in vivo to allow AAV gene therapy applications. DIMMERs should thus allow routine research and biotechnology applications to operate with greater precision, control, and predictability.

2.3 Results

2.3.1 A minimal model shows that dosage compensation requires linear sensitivity to miRNA

To guide the design of DIMMERs, we first developed a minimal model of a miRNA-based IFFL circuit. This model makes several assumptions: (1) Primary miRNA (pri-miRNA) and target mRNA are each transcribed constitutively at a fixed ratio of rates, and in direct proportion to gene dosage. (2) There is a constant total rate, per gene copy, of RISC complex production, reflecting the combined process of pri-miRNA transcription, post-transcriptional processing, and binding to Argonaute proteins [12]. (3) RISC and its target mRNA bind reversibly to form a RISC-mRNA complex. Finally, (4) formation of this complex leads to degradation of the bound mRNA.

In certain regimes, this model exhibits dosage-invariant expression profiles, in which target protein expression levels increase linearly and then asymptotically approach a dosage-independent limiting expression level (**Figure 2.1D**). Several parameters modulate the limiting expression level while preserving asymptotic dosage compensation. These include the binding and unbinding rates of mRNA to miRNA, and the catalytic rate of mRNA degradation. Accessing the dosage compensation regime requires that mRNA levels be linearly sensitive to RISC concentration. Below, a more detailed model in which the total amounts of free and bound mRNA were considered, revealed that miRNA-dependent catalytic degradation rates must exceed a minimal value for the total mRNA to show dosage invariance (**Data S1**). Together, these results suggest conditions in which miRNA-based IFFL circuits could produce gene dosage-invariant expression.

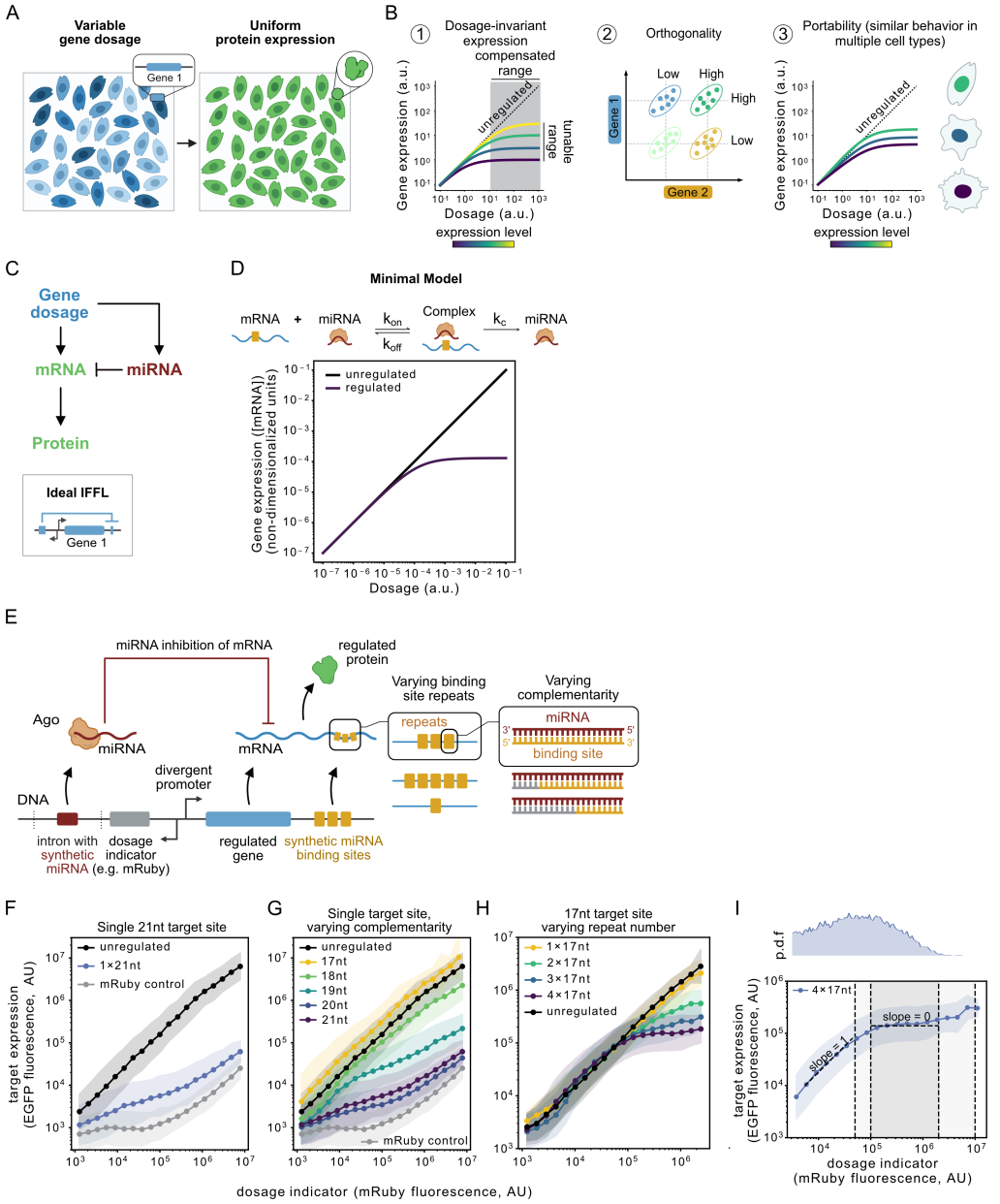

Figure 2.1: miRNA incoherent feedforward circuits enable dosage-invariant gene expression

(A) An ideal gene expression system generates uniform protein expression (right) levels despite variable gene dosage delivered (left).

(B) An ideal system would enable tunable (①), orthogonal (②) control of the target and operate in multiple cell contexts (③). (②) shows that two genes can be regulated by two orthogonal circuit designs with tunable expressions. Ellipses, the majority of protein expression profiles. Dashed lines, mean expressions.

(C) The architecture of the incoherent feedforward loop (IFFL), top panel. Simplified miRNA-based IFFL circuit configuration, bottom panel.

Figure 2.1: (continued)
(D) The minimal model shows that miRNA-mediated regulation enables dosage compensation compared to the unregulated control. Upper panel, miRNA-mediated regulation reactions. See also **Data S2**.
(E) The circuit configuration of the miRNA-mediated IFFL.
(F) A single, fully complementary miR-L site (1×21nt) shows strong repression compared to the unregulated but does not achieve dosage compensation. The mRuby control shows the bleed-through signal from mRuby3 in the EGFP channel. Cells are gated and binned by mRuby3 intensities. Each dot corresponds to the geometric mean fluorescence intensity of mRuby3 bin breaks and median fluorescence intensity of EGFP in the bin. Shaded regions, geometric standard deviation.
(G) Reducing the complementarity (pairing length starting from the seed region) of the single-site target weakens the repression but does not achieve dosage compensation.
(H) Multimerizing the 17nt site restores the repression and achieves dosage compensation at four repeats.
(I) A zoom-in of the 4×17nt in **(H)**. p.d.f, probability distribution function of mRuby3 fluorescence intensity. The gray and light gray rectangles indicate the dosage range where the fluorescence intensity of EGFP does not change or change by 4-fold, respectively. The dashed lines with slope=0 or slope=1 indicate the linear dosage dependence at low dosage or dosage independence at high dosage, respectively.
(F)-(H), the x-axis and the y-axis show the dosage indicator expression and the target expression, respectively, both in arbitrary units (a.u.).

2.3.2 Multimerization of weak target sites provides dosage compensation

Based on these results, we designed an initial set of regulatory circuits. Briefly, a synthetic miRNA and a target mRNA were placed in opposite orientations relative to a central divergent promoter [7] (**Figure 2.1E**). In one orientation, a previously characterized synthetic miRNA (miR-L, based on a Renilla luciferase sequence [13]) was co-expressed with the fluorescent protein mRuby3, serving as a dosage indicator. The miRNA expression cassette included the miR-E backbone for pri-miRNA expression [13] and was incorporated within a synthetic intron [14]. In the opposite orientation, we inserted a constitutively expressed EGFP target gene, with a single fully complementary 21-nt miRNA target site in the 3'-UTR. This format allowed analysis of multiple miRNA and target site configurations, and independent modulation of miRNA and target gene expression levels. We also systematically analyzed a broad variety of other architectures (**Data S2**).

To quantitatively measure the behavior of the circuit, we transfected U2OS cells with each circuit construct and analyzed expression by flow cytometry 48h later. We then plotted target EGFP expression versus gene dosage, indicated by mRuby3

fluorescence (**Figure 2.1F**). Compared to an unregulated control with no miRNA target site, the circuit strongly reduced target EGFP expression by 1–2 orders of magnitude. However, it failed to achieve dosage compensation (**Figure 2.1F**). We also analyzed similar circuits in which the complementary region was systematically reduced in single nucleotide increments from 21nt to 17nt. Constructs with 19 or 20nt sites retained repression but failed to produce dosage compensation, while shorter sequences lost repression altogether (**Figure 2.1G**). Thus, single target sites with varying levels of complementarity did not provide dosage compensation.

Native miRNAs are known to use much shorter complementary regions, including central mismatches (bulges) [12], and multivalent interactions mediated by TNRC6 scaffold proteins [11, 15]. Therefore, we next considered designs with reduced complementarity and multimerized target binding sites. Tandem repeats of two to four copies of the 17nt target site, which was inactive in a single copy, progressively increased regulation (**Figure 2.1H**). More importantly, they successfully reduced dosage sensitivity, particularly at higher expression levels (**Figure 2.1H, I**). Thus, 4 tandem 17nt sites yielded only a ~4-fold change in expression over a ~200-fold range of dosages (**Figure 2.1I**). We also tested tandem repeats of two to four copies of the 18nt and 19nt target binding sites. Interestingly, the 18nt target site showed dosage compensation behavior at 3 copies, while the 19nt target site exhibited dosage compensation with only 2 copies (**Figure S2.1A, B**). "Bulged" target sites that provided no regulation individually nevertheless exhibited strong regulation and limited dosage compensation when multimerized (**Figure S2.1C, D**). These results indicate that multiple tandem copies of individually weak target sites can achieve dosage compensation over broad dosage regimes, at varying setpoints.

2.3.3 TNRC6 and Ago2 play key roles in regulation of multimerized 17nt targets

Why do multimerized weak sites produce better dosage compensation than stronger individual target sites? We reasoned that the TNRC6 scaffold protein could facilitate inhibition of multimerized weak target sites. To test whether regulation of multimeric weak sites requires TNRC6, we expressed T6B, a previously identified TNRC6B protein fragment that competitively inhibits TNRC6 activity (**Figure 2.2A**) [16]. T6B abolished regulation by a co-transfected 4×17nt DIMMER (**Figure 2.2B**), but had little effect on the single fully complementary 21nt construct (**Figure 2.2C**). As expected, negative controls using a T6B variant lacking the Ago2-binding domain failed to abolish regulation (**Figure S2.1E**). Interestingly, overexpression

of wild-type Ago2 exhibited similar effects as ectopic T6B expression, nearly eliminating regulation in the 4×17nt case (**Figure 2.2D**), without affecting the 1×21nt configuration (**Figure 2.2E**), possibly due to sequestration of TNRC6. Thus, TNRC6 activity is required for regulation of the multimeric 4×17nt target, but not the single fully complementary site.

Although Ago2 predominantly directs repression through a slicing-independent mechanism, slicing occurs in at least a dozen complementary targets [17], and is required for maturation of two erythroid miRNAs that undergo non-canonical biogenesis [18]. To understand whether slicing activity is required for dosage compensation, we ectopically expressed a dominant negative Ago2 mutant (D669A) lacking slicing activity. This perturbation reduced regulation of both fully complementary and multimerized partially complementary targets (**Figure 2.2F, G**), suggesting that Ago2 slicer activity is required for repression of both targets. Further, to test for a role in cleavage of passenger strands, we incorporated a mismatch at 10-11nt in the original miRNA design (**Figure 2.2H**) in order to eliminate the requirement for Ago2 cleavage in miRNA maturation [18]. Inclusion of the mismatch slightly enhanced regulation of both the 4×17nt and 1×21nt (**Figure S2.1K, L**), possibly by elevating the efficiency of miRNA biogenesis. Although passenger strand removal is not required in this case, dominant negative Ago2 expression nevertheless largely eliminated regulation of both targets, consistent with a role for Ago2-dependent slicing in target repression (**Figure 2.2I-J**). By contrast, wild-type Ago2 overexpression had a partial impact on the 4×17nt target but no impact on the 1×21nt target (**Figure 2.2I-J**). We also analyzed other target sequences. The 4×18nt and 4×19nt targets responded similarly to T6B and Ago2^{DN}, with Ago2^{DN} showing a slightly stronger de-repression effect. By contrast, wild-type Ago2 had weaker or no suppressive effect on these stronger target sites (**Figure 2.2K-L**), possibly because their longer complementary regions were less sensitive to indirect sequestration of TNRC6. Taken together, these results suggest that TNRC6 is required for repression of multimerized partially complementary targets but not fully complementary targets, and that Ago2-dependent slicing is required for repression of partially and fully complementary targets.

Taken together, these results suggest a potential explanation for why dosage compensation requires multimerized weak binding sites and TNRC6 (**Figure 2.2M**). Briefly, fully complementary binding sites are known to produce much more efficient slicing than partially complementary sites [19]. The fully complementary

21nt target could therefore produce dosage compensation at dosages and setpoints comparable to or lower than detection limits in the flow cytometry experiments. At the same time, the higher affinity of fully complementary binding sites would shift the "tail" of elevated expression due to bound-but-not-yet-degraded complexes to lower dosages. These combined effects could together explain the shape of the single 21nt fully complementary construct (**Figure 2.1F**). By contrast, constructs with multimerized weak binding sites (e.g. 4×17nt) would engage at higher gene dosages due to the lower intrinsic RNA affinity. This would lead to elevated setpoints and shift the "tail" of non-compensated expression to correspondingly higher dosages (**Figure 2.1I, Data S1**). Efficient mRNA degradation could be achieved by enhancing the intrinsically weaker Ago2-dependent slicing activity through stabilization of Ago2-mRNA complexes by TNRC6 and through TNRC6-dependent deadenylation, as described previously.

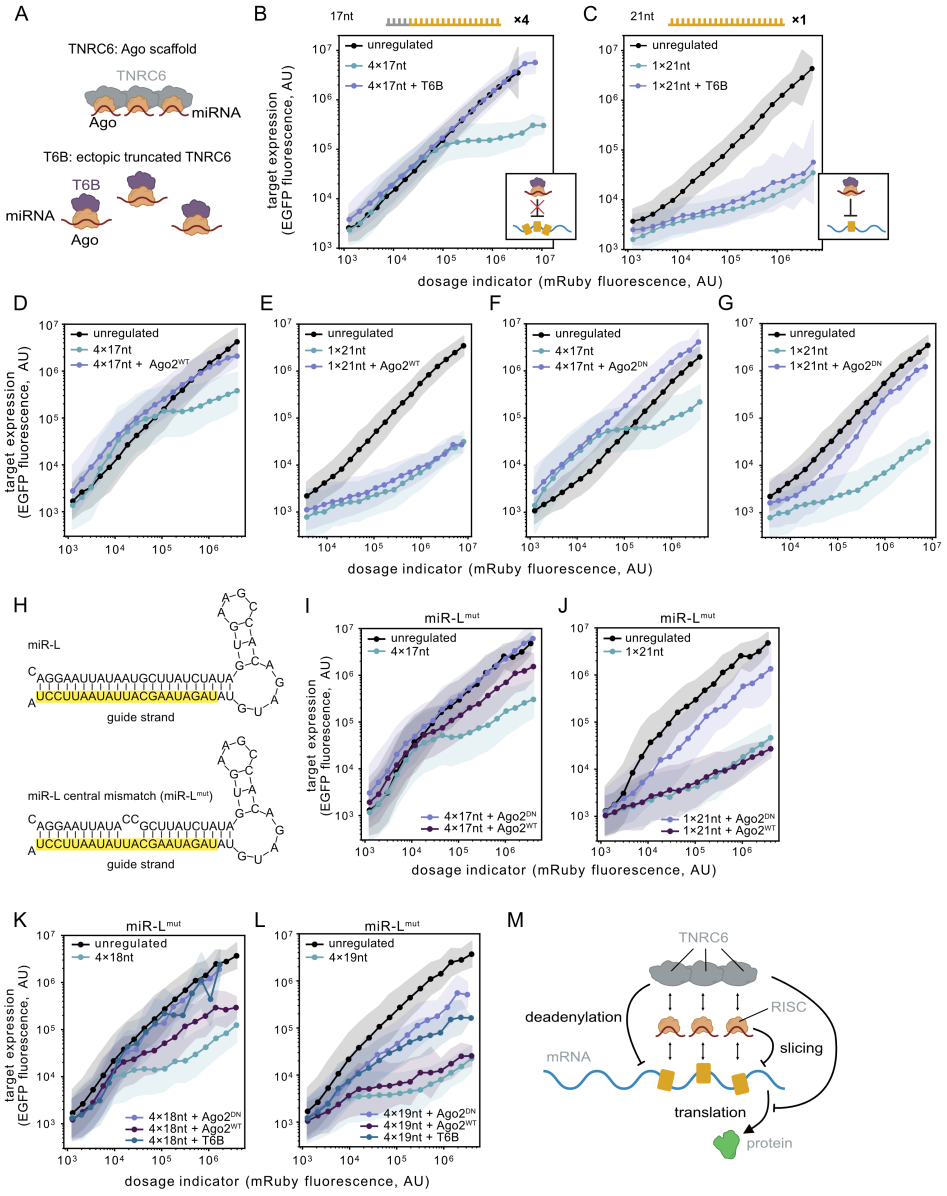

Figure 2.2: TNRC6 and Ago2 play key roles in regulation of multimerized weak targets

(A) T6B peptide competitively inhibits TNRC6-dependent regulation (schematic).
(B, C) T6B expression suppresses miR-L-mediated regulation of 4×17nt construct **(B)** but not 1×21nt **(C)** construct.
(D, E) Ectopic wild type Ago2 ($Ago2^{WT}$) suppresses inhibition of 4×17nt **(D)** but not 1×21nt **(E)** construct.
(F, G) Dominant negative Ago2 ($Ago2^{DN}$) suppresses regulation of both 4×17nt **(F)** and 1×21nt **(G)** constructs.

Figure 2.2: (continued)
(H) Design of synthetic miR-L cassettes without (top) or with (bottom) central mismatches in the passenger strand.
(I, J) Ectopic Ago2^{WT} or Ago2^{DN} modulate miR-L^{mut}-mediated regulation of 4×17nt **(I)** and 1×21nt **(J)** constructs.
(K, L) The effect of T6B/Ago2^{WT}/Ago2^{DN} overexpression in the miR-L^{mut}-mediated regulation for 4×18nt **(K)** and 4×19nt **(L)** constructs.
(M) TNRC6 and Ago2 play important roles in dosage compensation (schematic).
In **(B)-(C)**, **(D)-(G)**, and **(I)-(L)**, in unperturbed groups, cells were co-transfected with filler plasmids to match the total transfection amount.
See also **Figure S2.1**.

2.3.4 Dosage-invariant expression levels can be tuned

Having demonstrated dosage invariant circuit designs and gained insight into their design from the model, we next sought to identify experimentally tunable parameters that modulate expression level setpoints (**Figure 2.1E**). We first examined varying the length of target complementarity from 8 to 21nt, while maintaining 4 tandem repeats (**Figure 2.3A, Data S2**). Repression was modest at 8nt, diminished with increased complementarity in the central region, and then strengthened again as more complementarity was added after the central region (**Figure 2.3A,** left panel). These results are consistent with previous observations that miRNA inhibition does not increase monotonically with complementarity [20, 21, 22]. For the miR-L target site, repression was most sensitive at 16-20 nucleotides of complementarity (**Figure 2.3A**). Three designs—4×17, 4×18, and 4×19— achieved dosage invariant expression, but did so at distinct setpoints spanning more than an order of magnitude in saturating expression level (**Figure 2.3B**). These results indicate that site length can tune setpoint.

We next varied the number of copies of the miRNA expression cassettes in the synthetic intron, effectively modulating the stoichiometric ratio of miRNA to mRNA (**Figure 2.3C**). Compared to a single copy, two or three copies of the miRNA reduced expression by 2-fold and 3-fold, respectively, while preserving dosage compensation, providing a means of fine-tuning expression control.

We also compared different promoters to vary transcription of the miRNA and target cassettes. For the miRNA the weaker PGK promoter allowed ~3.5-fold more target gene expression at a given dosage level compared to the stronger EF1α, but nevertheless preserved dosage compensation (**Figure 2.3D**). Finally, we incorporated a tetracycline-responsive system (**Figure 2.3E**). This allowed setpoint tuning over ~2

orders of magnitude while preserving dosage compensation (**Figure 2.3E, Data S2**). Taken together, these results demonstrate that dosage invariance can be preserved while allowing multiple mechanisms of coarse (site length, target transcription rate) and fine (miRNA promoter and copy number) setpoint tuning.

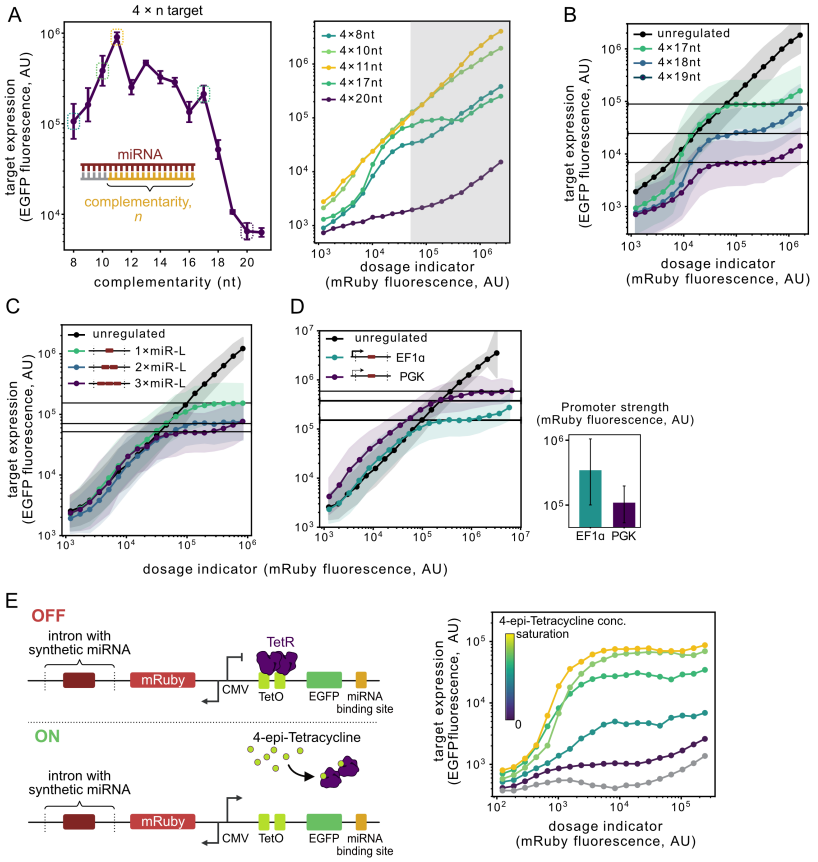

Figure 2.3: Dosage-invariant expression levels can be tuned
(A) 4×miR-L target expression levels exhibit non-monotonic complementarity dependence. Expression (y-axis) represents median EGFP fluorescence intensity in cells with mRuby3 fluorescence intensity > $5{\times}10^4$, see also the shaded region in the right panel). Data represents 3 biological replicates. Right panel, dosage response curves of selected constructs (dashed boxes in the left panel).
(B) Complementarity of 4×17-19nt monotonically regulates setpoint (black horizontal line).
(C) Number of miR-L cassettes fine-tunes setpoint (black horizontal line).
(D) miR-L promoter strength regulates setpoint (black horizontal line). Bars indicate mRuby3 median fluorescence intensities of transfected cells. Data represents 3 biological replicates. In C and D, the target is 4×17nt.

Figure 2.3: (continued)
(E) Tet system allows tunable dosage compensation. Left panel, design of the inducible DIMMER. A CMV promoter harboring two downstream TetO sites drives 4-epi-Tc inducible expression in the TRex cell line. Right panel, the inducible 4×19nt DIMMER dosage response curve. Drug concentrations from purple to yellow were 0, 10, 33.3, 100, 333.3 ng/mL. Gray curve, mRuby-only transfection control.
See also **Data S2**.

2.3.5 Orthogonal dosage compensation circuits allow independent control of target genes

Engineered genetic systems increasingly require multiple genes and transcripts, provoking a need for multiple independent DIMMER systems based on orthogonal synthetic miRNA-target site pairs, which we term *synmiRs*. To design synmiRs, we started by generating five random miRNA sequences, labeled synmiR1-5. Each sequence contained A at position 1 in the miRNA, a cognate U at the 3' end of a single target site and 25% GC content, similar to the structure of miR-L. We used an "open loop" system to analyze their behavior, allowing independent control of miRNA expression and measurement of its effect on a target miRNA reporter gene (**Figure 2.4A**). For a single fully complementary miR-L site, inhibition increased with miRNA dosage (**Figure 2.4B**), consistent with the earlier closed loop results (**Figure 2.1F**). synmiRs 1,4, and 5 repressed by at least an order of magnitude relative to a control lacking the miRNA (**Figure 2.4D, Data S2**). By contrast, synmiRs 2 and 3 achieved weaker repression, possibly due to subsequences containing two or more A/T pairs in the extensive region, which could destabilize the miRNA [23, 24] (**Figure S2.2A**). Consistent with this hypothesis, an A to G substitution at position 20 in synmiR-2 or at position 19 in synmiR-3 restored miRNA inhibition of target gene expression (**Figure 2.4D, Data S2**).

Based on these results, we formulated an empirical algorithm for synmiR design. Briefly, we generate random 21nt candidate miRNA sequences in which the mature miRNA (1) includes a 5'-U, based on known requirements for miRNA loading [12, 25], and, (2) includes 5–8 G or C nucleotides, with 1–4 of them in the seed region, and 1–2 in the extensive region (**Figure 2.4C**). We synthesized five candidate sequences (synmiRs 6-10) based on this simple algorithm and analyzed their open loop behavior (**Figure 2.4D**, **Figure S2.2B**). All five sequences generated strong repression as single sites, comparable to that of miR-L. Altogether, these results produced ten miRNA sequences capable of strong repression in their fully complementary form

(**Figure 2.4D, Data S2**). Pairing each of these ten miRNAs with all ten of the target sequences in the open loop system revealed strong orthogonality in regulation, as desired (**Figure 2.4E, Figure S2.2C**).

Using these synmiR sequences, we developed a set of ten orthogonal dosage compensation circuits, using the framework in **Figure 2.1E**. Each of these sequences could produce dosage compensation in some configurations and dosage regimes. For example, with synmiR-4 and synmiR-5, the 4×17nt configuration produced a strong inhibition profile more similar to the fully complementary target (**Figure S2.2D**). We reasoned that this could reflect higher GC content in the seed and supplementary regions of these two miRNAs compared to miR-L, allowing shorter 8 or 9 nt target sites, present in 4-8 repeats, to produce dosage compensating designs (**Figure 2.4F**). In a similar way, we identified dosage compensation regimes for the other sequences (**Figure 2.4F**). Eventually, we identified nine DIMMER circuits exhibiting different levels of dosage compensation, of which five showed at least a 30-fold dosage-invariant range (**Figure 2.4F, Data S2**). Further, expression levels in these circuits were sensitive to the number and complementarity of target sites, allowing setpoint tuning (**Figure S2.3A**). These designs relied TNRC6 for dosage compensation (**Figure S2.3B**). The exception was synmiR-2 8×8nt, which maintained strong repression even under the T6B perturbation (**Figure S2.3B**). Taken together, these results provide a toolkit of dosage compensating systems and, more generally, suggest that it should be relatively feasible to engineer many additional dosage compensating systems with varying expression setpoints.

With multiple dosage compensation systems, it should be possible to independently specify the expression of multiple target genes in the same cells (**Figure 2.1B, panel 2**). To test this possibility, we constructed a second set of dosage compensation expression systems using distinct fluorescent reporters (**Figure 2.4G**). We transfected cells with pairs of systems that had different regulatory setpoints, and analyzed the resulting expression profiles of the two regulated target genes (**Figure 2.4H, Figure S2.2E**). Altogether, we analyzed four pairs of systems. Each produced a distinct two-dimensional expression distribution based on the setpoints for the two reporters (**Figure 2.4H**). By contrast, the unregulated group showed higher setpoints, and broader distributions of both reporters. The dosage indicators also exhibited the same distributions among all groups. Interestingly, the circuits allowed precise control of the stoichiometry of the regulated proteins (**Figure S2.2E**). The engineered dosage compensation systems thus make it possible to specify two-dimensional ex-

pression distributions, and suggest that control of higher dimensional distributions of more genes should also be accessible.

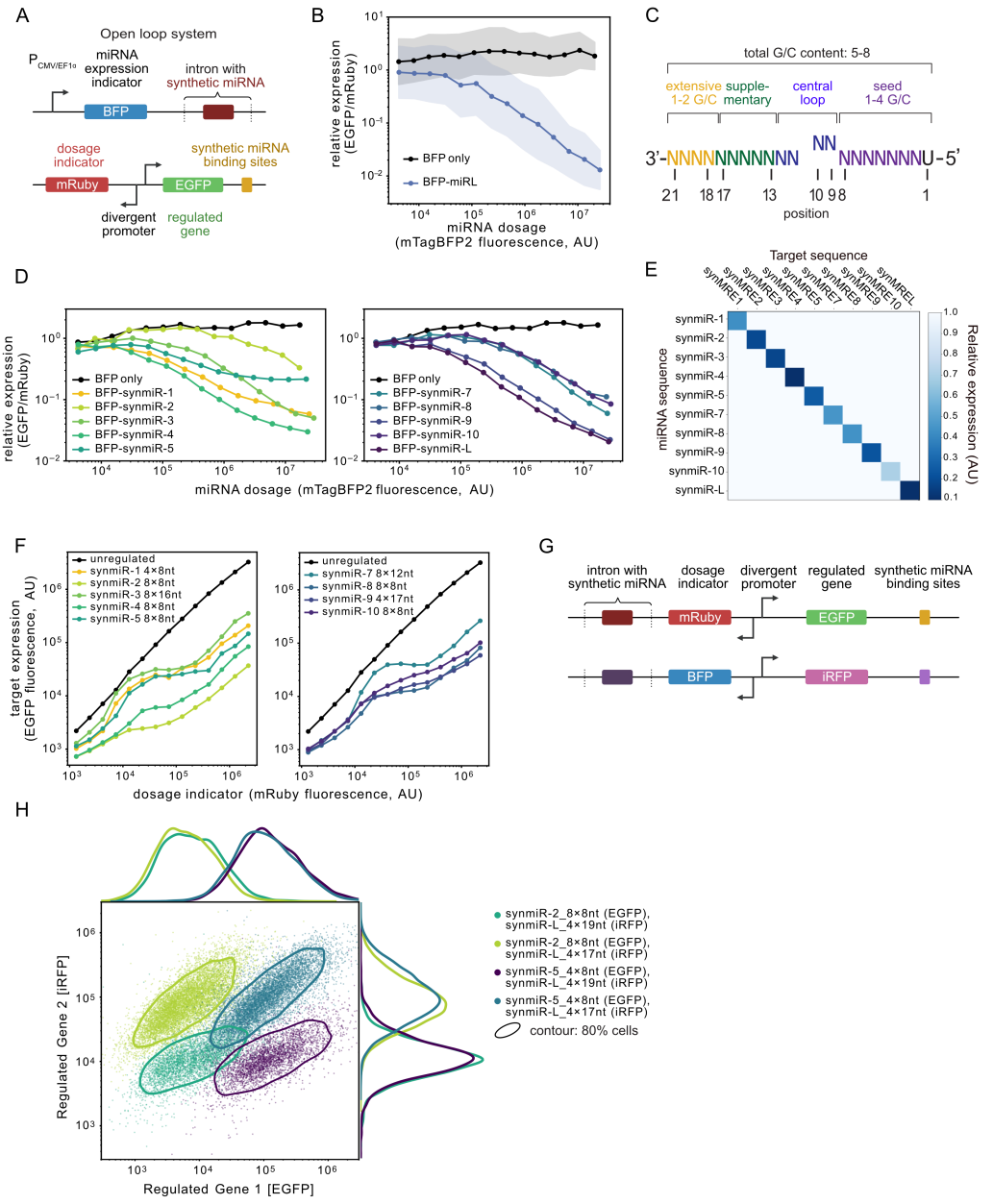

Figure 2.4: Orthogonal dosage compensation circuits allow independent control of target genes

(A) Open loop system allows simultaneous readout of miRNA expression (mTagBFP2) and its effect on target gene expression (EGFP), at different target gene dosages (mRuby3).

Figure 2.4: (continued)
(B) Open loop circuits reveal quantitative miRNA input-output relationship. Cells were co-transfected with the 1×21nt target, either with the miR-L construct shown in **(A)**, or mTagBFP2 as a control. Relative expression, median ratio of EGFP fluorescence intensity to mRuby3 fluorescence intensity in each mTagBFP2 bin. Dots and shaded regions are as described in **Figure 2.1F**.
(C) synmiRs use a simple empirical design algorithm.
(D) All synmiRs can regulate cognate single 21nt targets.
(E) Ten synmiRs function orthogonally. Cells were co-transfected with each combination of target and miRNA for open loop analysis. Relative expression levels were quantified as described in **(B)** and normalized by maximal expression of each construct across synmiRs.
(F) synmiRs can generate dosage compensated regulation.
(G) Two DIMMER constructs can be analyzed in the same cells using distinct fluorescent proteins (schematic).
(H) The double DIMMER system allows independent regulation of two genes in four distinct expression configurations. Cells were poly-transfected with the double DIMMER reporter system shown in **(G)**. Data were gated by mRuby3 and mTagBFP2 intensities. Solid contour lines indicate regions containing 80% of cells. Profile plots show the distributions of the corresponding fluorescent proteins.
See also **Figure S2.2-S2.3**, and **Data S2**.

2.3.6 Dosage compensation systems are portable and minimally perturbative

An ideal dosage compensation system would be portable, able to operate similarly across different cell types, function in both transient transfection and genomic integration, and minimally perturb the host cell. To examine these features, we transiently transfected several circuit variants, including the 4×17nt miR-L system (**Figure 2.1H**), in four mammalian cell lines: U2OS [26], CHO-K1 [27], HEK293 [28], and N2A [29]. In each cell line, we observed strong and qualitatively similar dosage compensation (**Figure 2.5A**). Cell lines varied in the threshold dosage at which expression saturated (**Figure 2.5A**, gray vertical line), and in the saturating expression level (**Figure 2.5A**, gray horizontal line), as measured in arbitrary fluorescence units. However, the ratio of these values was conserved (**Figure 2.5B**). We obtained similar results for other circuits as well, including synmiR-4, with 8 repeats of a 9nt target site, as well as both synmiR-L and synmiR-5, each with 8 repeats of an 8nt target site (**Figure S2.4A**). Again, the ratio of the saturating expression level to the threshold dosage was similar, for each construct, across cell lines (**Figure S2.4B**). This suggests a model in which the miRNA circuit functions equivalently in different cell types, but protein expression strengths vary, possibly

due to differences in translational capacity or basal protein degradation rates [30, 31, 32, 33]. Together, these results indicate that the dosage compensation circuits can function across different cell contexts.

Stable cell lines are important in research as well as applications like cell therapy. To find out whether dosage compensation circuits could also function in a stable integration context, we used PiggyBac transposition together with the iON system that allows expression only from constructs that have successfully integrated in the genome and undergone site-specific recombination [34]. We then selected mono-clones, and analyzed reporter expression by flow cytometry (**Figure 2.5C**). Integration copy numbers varied among clones by over two orders of magnitude, as indicated by mRuby3 fluorescence intensity (**Figure 2.5D left panel**, x-axis). Nevertheless, the cargo EGFP expression remained nearly constant (**Figure 2.5D left panel,** y-axis). By contrast, the unregulated mono-clones exhibited an apparent correlation between the integrated copy numbers and the EGFP expression (**Figure 2.5D right panel**). Thus, dosage compensation circuits function in stable integration settings.

The expression of synthetic miRNAs could in principle perturb endogenous gene expression. To identify such effects, we performed bulk RNA sequencing on cells transfected with miR-L and each of the 9 orthogonal synmiRs, and compared them to a negative control transfection of a BFP expression vector. Only a few genes were significantly up- or down-regulated by the miRNA (**Figure 2.5E**). These were enriched for heat shock proteins such as HSPA6. Critically, the gene sets up-regulated by different miRNAs exhibited strong overlap (**Figure 2.5F**, **Data S2**). Thus, for the synmiRs described here, off-target regulation appears to only reflect non-specific effects of miRNA expression, rather than sequence-specific perturbations.

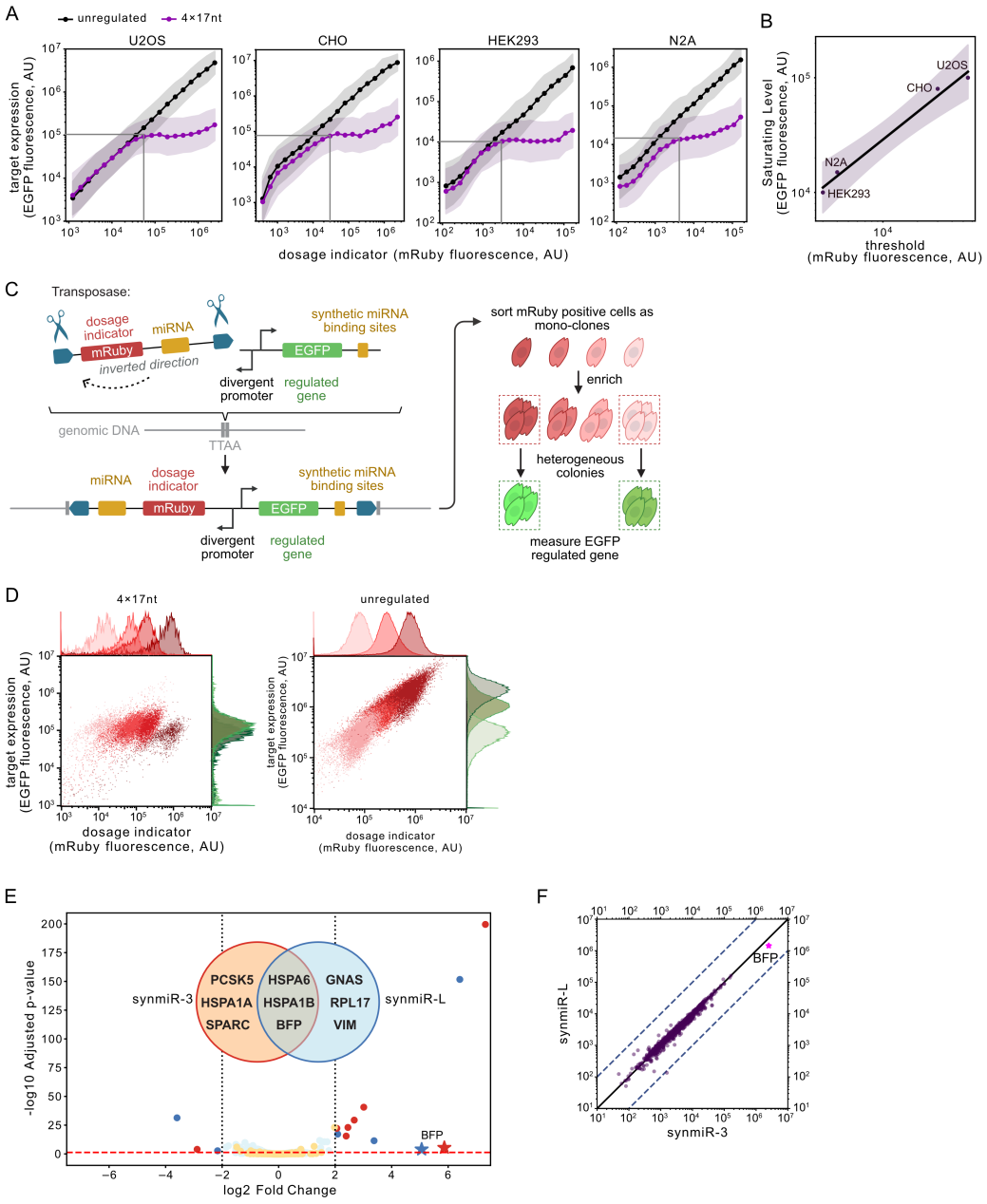

Figure 2.5: Dosage compensation systems are portable and minimally perturbative

(A) DIMMERs function in multiple cell lines. Gray vertical lines, minimal dosage for the compensation regime (threshold); horizontal lines, setpoint level.

(B) The threshold and the setpoint level co-vary across cell lines. Black solid line, data linear fit in the logarithm space (**STAR Methods**).

(C) iON transposition genomically integrates DIMMER circuits.

Figure 2.5: (continued)
(D) DIMMER circuits enable uniform protein expression (y-axis) across a range of stable integrations (x-axis). Each color shows the 4×17nt DIMMER (left panel) compared with the unregulated construct (right panel) in PiggyBac-integrated monoclonal U2OS cells. Profile histograms show mRuby3 and EGFP distributions of each mono-clone.
(E) DIMMER circuits show limited perturbation to the endogenous transcriptome. Bulk RNAseq volcano plots show cells transfected with either synmiR-3 (orange and red dots) or synmiR-L (light blue and dark blue dots), compared to mTagBFP2-only transfected cells. Vertical dashed lines, $|\log_2(\text{fold change})| = 2$. Horizontal red dashed line, adjusted p-value = 0.05. The Venn diagram shows differentially expressed genes between synmiR-3 and synmiR-L.
(F) Global mRNA expression levels in normalized transcripts per million (TPM) of miR-L expressing cells are plotted against those of synmiR-3 expressing cells. Solid line, equal expression in both samples. Dashed lines, 10-fold expression differences. See also **Figure S2.4** and **Data S2**.

2.3.7 Dosage compensation enhances biological imaging

Single-molecule imaging approaches such as DNA-PAINT (Point Accumulation for Imaging in Nanoscale Topography) often rely on ectopic expression of tagged proteins [35], which can dramatically exceed endogenous expression, distorting subcellular localization patterns. DIMMER circuits could potentially address this issue by limiting ectopic protein expression.

To test this, we transfected a DIMMER-regulated and unregulated EGFR-mEGFP membrane marker fusion protein expression-constructs into CHO-K1 cells, which lack endogenous EGFR expression [36] (**Figure 2.6A**). The DIMMER circuits successfully reduced expression, as measured by receptor density relative to an unregulated control, by at least 10-fold, as estimated from flow cytometry and confocal imaging (**Figure S2.5A-C**). To quantify this reduction more directly, we imaged transfected cells 48h post-transfection using the DNA-PAINT method, based on DNA-conjugated anti-GFP nanobodies targeting the intracellular mEGFP tag of the EGFR (**Figure 2.6B-D**). Circuit-regulated receptor mean densities were $\tilde{1}$5-47 times lower than unregulated receptors (**Figure 2.6E, Figure S2.5D**). Further, expression distributions were narrower for the regulated plasmids compared to the unregulated ones (**Figure 2.6E**). Critically, DIMMER reduced expression more homogeneously than possible by simply reducing the concentration of unregulated plasmids in transfections (**Figure 2.6F**, **Figure S2.5D**). DIMMER-regulated expression with the 4×17nt circuit achieved EGFR expression levels comparable to

but lower than those of wild-type U2OS cells. Thus, DIMMER circuits allow homogeneous reduction in ectopic expression to endogenous levels.

CRISPR imaging methods enable analysis of specific genomic loci in cell nuclei, but can be limited by high background fluorescence due to basal expression of dCas9 fluorescent protein fusions [37, 38]. DIMMER circuits also improved dCas9-EGFP imaging of telomeres. We transfected dCas9-EGFP with or without DIMMER circuits, along with a gRNA targeting repetitive telomeric sequences [38, 39, 40] (**Figure 2.6G**). The circuits reduced dCas9-EGFP expression, and its dosage sensitivity (**Figure S2.6A**). Critically, in the unregulated system, dCas9-EGFP formed bright aggregations in the nucleolus, but seldom labeled the telomeres, consistent with previous observations [41] (**Figure 2.6H left panel**). By contrast, the 4×17nt circuit restricted most fluorescence to puncta (dots), consistent with telomeric labeling [38, 39, 40], and reduced labeling of the nucleolus (**Figure 2.6H middle panel**). Further, the stronger 4×19nt circuit removed nearly all labeling in the nucleolus, while maintaining apparent telomere labeling (**Figure 2.6H right panel**). DIMMER circuits also substantially improved signal-to-background ratio as well as contrast for individual dots (**Figure S2.6B**). Taken together, these results demonstrate that dosage compensation circuits can improve imaging of proteins and subcellular structures.

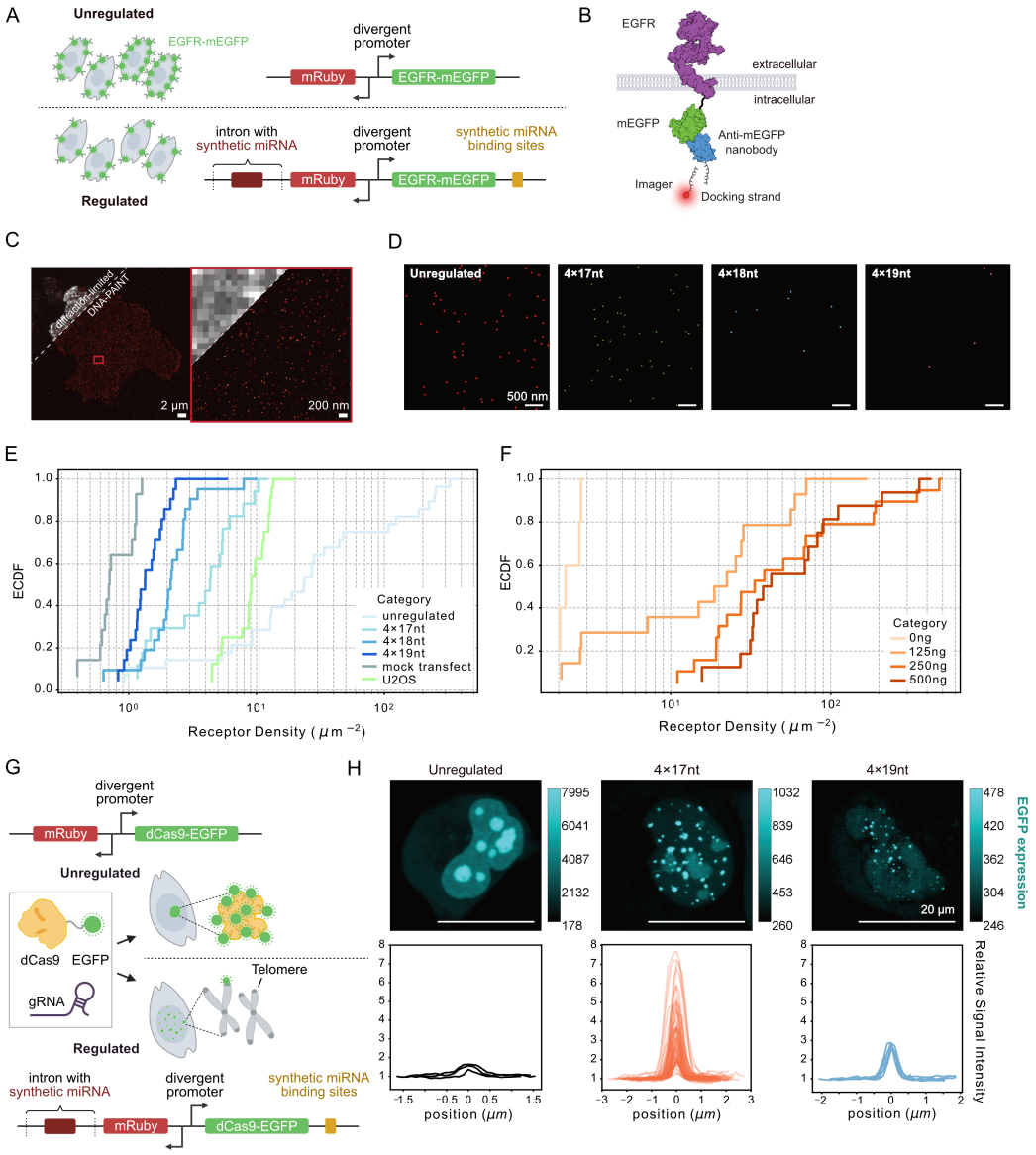

Figure 2.6: Dosage compensation circuits enhance biological imaging
(A) EGFR fusions to mEGFP were expressed with (lower panel) or without (upper) regulation.
(B) EGFR (adapted from PDB 7SYE) is linked to an mEGFP labeled with a DNA-conjugated anti-GFP nanobody (PDB 6LR7).
(C) Diffraction-limited and DNA-PAINT imaging of CHO-K1 cells expressing EGFR-mEGFP. Right panel, zoom-in of the boxed area.
(D) Representative single-molecule images of EGFR-mEGFP. Images show (L-R) unregulated or regulated with DIMMER variants.

Figure 2.6: (continued)
(E) DIMMER circuits homogeneously reduce expression levels to near-background levels. The empirical cumulative distribution function (ECDF) of the receptor density shows CHO-K1 cells transfected with EGFR-mEGFP with or without the DIMMER module (blue curves), and the endogenous EGFR density in U2OS cells (green curve), along with mock-transfected CHO-K1 cells, measured by DNA-PAINT. Receptor densities: unregulated: 79.0 ± 42.4 μm^{-2}, 4×17nt: 5.1 ± 1.7 μm^{-2}, 4×18nt: 2.8 ± 1.0 μm^{-2}, 4×19nt: 1.7 ± 0.5 μm^{-2}, mock transfection: 0.9 ± 0.2 μm^{-2}, U2OS: 9.8 ± 1.5 μm^{-2} (mean ± 95% confidence interval).
(F) Reducing unregulated plasmid concentration does not provide a low and homogeneous receptor density, measured by DNA-PAINT. Receptor densities: 0 ng: 2.5 ± 0.4 μm^{-2}, 125 ng: 36.0 ± 25.0 μm^{-2}, 250 ng: 115.7 ± 74.4 μm^{-2}, 500 ng: 104.1 ± 64.5 μm^{-2}(mean ± 95% confidence interval).
See also **Figure S2.5**.
(G) dCas9 can be used to image telomeres (schematic). Telomere-targeting gRNA is driven by a U6 promoter and expressed from the dCas9-EGFP vector.
(H) DIMMER circuits enable telomere imaging by suppressing background. Top row: representative images of dCas9–EGFP telomere labeling in the unregulated construct and in two DIMMER designs (4×17nt and 4×19nt). Color bars, measured EGFP intensities; scale bars, 20 µm. Bottom row: line-scan profiles of relative signal intensity for individual puncta. In the unregulated condition, high background obscures isolated dots.
See also **Figure S2.6**.

2.3.8 DIMMER reduces off-target RNA base editing and transcriptome stress

The adenine base editor ABEMax is a powerful gene-editing tool. However, it can induce undesired transcriptome-wide off-target A-to-I editing in HEK293 cells [42] and cause detrimental transcriptional responses [43], raising concerns for gene therapy applications. Engineering of the TadA and/or TadA* domain has partly reduced these issues [44, 45]. However, the rate of off-target RNA editing increases with the deaminase level [46]. We therefore asked whether DIMMER regulation could further reduce off-target RNA editing while preserving high levels of on-target DNA editing without additional protein engineering.

We investigated the impact of ABEMax regulation on both on-target DNA, and off-target RNA editing. To evaluate the on-target editing rate, we designed a guide RNA targeting the widely used "site 3" in HEK293 [45, 47]. We then co-transfected HEK293 cells with the guide RNA construct along with the ABEMax base editor, either unregulated or regulated with the DIMMER circuit module (**Figure 2.7A**). We initially chose the 4×18nt miR-L design since it provides an intermediate expression level setpoint. We harvested the cells 72h post-transfection, and extracted genomic

DNA and RNA for further measurements.

The regulated ABEMax base editor performed on-target editing at a similar, if modestly decreased, rate compared to the unregulated control (**Figure 2.7B**). To measure the off-target RNA editing rate, we selected three transcripts (*CTNNB1*, *RSL1D1*, *IP90*) known to exhibit high off-target editing due to similarity with native TadA tRNA substrates [44, 47]. The DIMMER circuit successfully decreased the mean A-to-I editing rate by 4 to 8.4-fold across the three transcripts (**Figure 2.7C**). It also reduced the number of detectable A-to-I conversions within the transcript (using a cutoff of 0.1% based on the sequencing results) (**Figure 2.7D**). At specific adenosines that are highly edited by the unregulated ABEMax, the circuit exhibited dramatically lower (around 10-fold) editing rates (**Figure 2.7E, Data S2**). We did not detect off-target genomic DNA editing on the predicted potential endogenous off-target sites with or without DIMMER. These results show that DIMMER regulation can improve the ratio of on- to off-target edits.

The DIMMER circuit also reduced perturbations to the transcriptome. We performed bulk RNA sequencing on cells treated with ABEMax with or without DIMMER regulation. DIMMER reduced transcriptome changes compared with unregulated ABEMax expression (**Figure 2.7F, Data S2**). More specifically, unregulated ABEMax exhibited $\tilde{2}$-fold more differentially regulated genes (adjusted $p<0.05$) compared to the regulated construct. Gene ontology annotation showed that heat shock proteins, DNA damage, and repair-associated proteins were perturbed in both unregulated and regulated ABEMax groups. However, the unregulated ABEMax group produced 89 differentially expressed genes in the 'nucleoplasm and nucleus' category, suggesting a more pervasive impact (**Figure 2.7F, Data S2, Table S2**). Together, these results indicate that DIMMER can improve the specificity of base editors by limiting their expression.

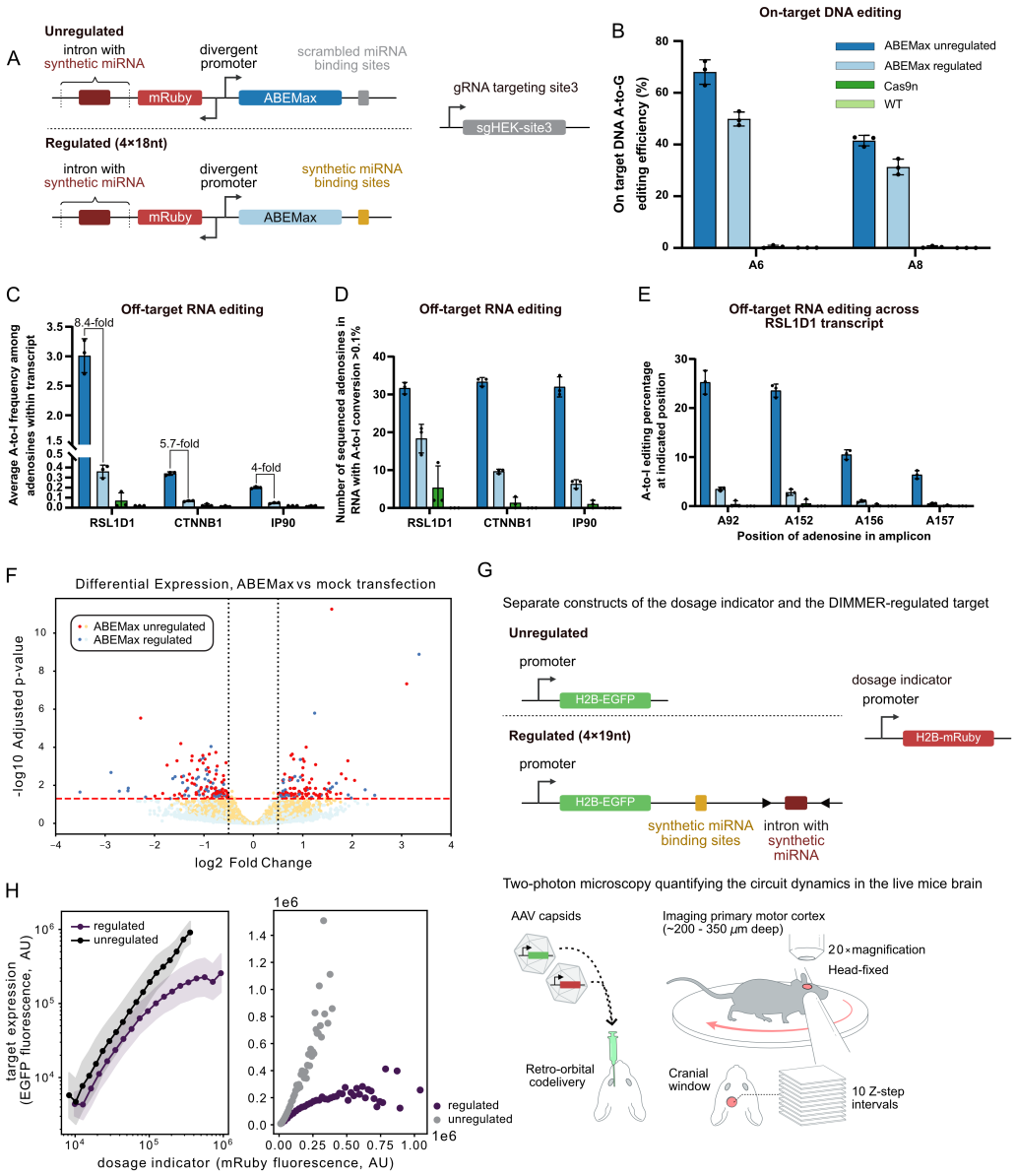

Figure 2.7: DIMMERs reduce ABEMax off-target RNA base editing and transcriptome stress, and operate in live mouse brains

(A) The ABEMax base editor was expressed with or without the miR-L-based DIMMER. The gRNA targeting HEK293 site 3 was expressed on a separate plasmid driven by the U6 promoter.

(B) DIMMER modestly reduces on-target A-to-G DNA editing efficiency, measured by the percentage of the A6 and A8 edited at HEK293 site 3.

(C-E) DIMMER strongly reduces off-target RNA editing. y-axis shows average off-target A-to-I RNA editing percentage **(C)** and the total number of sequenced adenosines in RNA with A-to-I conversion rate above 0.1% **(D)** for RSL1D1, CTNNB1, and IP90 across their transcripts, and off-target A-to-I RNA editing percentage of four specific A sites in the RSL1D1 transcript **(E)** (cf. **Data S2**).

Figure 2.7: (continued)
In **(B-E)**, each dot is a biological replicate. Error bars, standard deviation.
(F) DIMMER reduces off-target transcriptome perturbation. Volcano plots show bulk RNAseq of cells containing regulated (4×18nt) and unregulated ABEMax compared to mock transfection. Black dashed lines, $|\log_2(\text{fold change})| = 0.5$. Red dashed line, adjusted p-value = 0.05. Data represents three biological replicates.
(G) AAV-delivered constructs enable DIMMER-regulated protein expression in the mouse brain. Upper panel: the H2B-mRuby3 construct (dosage indicator) and the H2B-EGFP construct (target, either unregulated or regulated) driven by the same promoter were co-delivered to live mouse brains. Lower panel: two-photon microscopy experiments quantifying circuit dynamics in live mice brains. See details in **STAR Methods**.
(H) In vivo brain expression of the CaMKIIα promoter cohort measured by microscope. Data were pooled, binned and plotted based on mRuby3 fluorescence intensities on logarithmic (left) or linear (right) scales. See also **Figure S2.7**.

2.3.9 DIMMER circuits can operate in live mouse brains

Monogenic diseases of the central nervous system could potentially be cured by gene replacement therapies [48, 49] using recently developed adeno-associated virus (AAV) capsids that deliver genes across the blood-brain barrier [50, 51, 52]. However, many therapeutic genes are toxic or deleterious when overexpressed and expression must be maintained over long periods. These considerations underscore the need for controlled, durable cell-type specific in vivo expression of transgenes.

To test whether DIMMER circuits could function in vivo in mouse brain, we designed a set of 6 AAV vectors that express H2B-EGFP with or without the broadly dosage compensating 4×19nt DIMMER circuit (**Figure 2.7G upper panel, Figure S2.7A**), For in vivo expression, we compared three promoter systems with distinct cell-type specificities: the CaMKIIα promoter (CaMKIIα hereafter), which drives expression in excitatory projection neurons, particularly in the cortex and hippocampus [53, 54]; a hybrid CMV enhancer–MeCP2 promoter (MeP hereafter), which enables broad neuronal expression throughout the brain [55]; and a minimal β-globin promoter driven by the mDLX2 enhancer (mDLX2 hereafter), which selectively targets GABAergic interneurons in the forebrain, including the motor cortex [56]. As controls, we also delivered a corresponding set of unregulated constructs lacking the DIMMER circuit. All constructs were codelivered with the mRuby3 construct for internal normalization. We injected each pair of constructs into one cohort of mice, and analyzed expression of both fluorescent reporters over time by two-photon head-fixed imaging at single-cell resolution through a previously

described cranial implant [57] over the left motor cortex (**Figure 2.7G lower panel, STAR Methods**). This approach allowed us to monitor the fluorescent dynamics of EGFP and mRuby3 expression in individual cortical cells over timescales of 56-84 days.

To better understand the relationship between promoter strength, expression dynamics, and regulatory compensation, we analyzed circuit performance over time (**Figure S2.7B-E**). For all three promoters, fluorescence gradually accumulated over multiple weeks. Notably, the mDLX2 promoter, which is relatively weak and interneuron-specific, showed no evidence of dosage compensation across all timepoints (**Figure S2.7D**). Similarly, the MeP promoter, which drives moderate, panneuronal expression, shifted the distribution downward in regulated conditions but did not reach expression levels sufficient to trigger compensation, even at later time points (**Figure S7E**). In contrast, the strong CaMKIIα promoter exhibited dosage compensation at later stages (28–56 days). By 28 days, expression from the regulated CaMKIIα construct plateaued and diverged from the unregulated condition, consistent with engagement of post-transcriptional attenuation by the DIMMER circuit (**Figure 2.7H, S2.7B**). These findings suggest that dosage compensation can occur in the brain for some construct designs, that promoter strength and time-dependent expression dynamics are critical for entering the dosage compensation regime, and that weaker promoters may be unable to reach high enough expression levels to activate compensatory repression with this DIMMER variant (**Figure S2.7F**). In the future, increasing the number of miRNA copies or miRNA target sites could enable stronger dosage compensation with weaker promoters.

2.4 Discussion

Ectopic gene expression is a cornerstone of modern biology, and gene and cell therapy but precise control has remained elusive in most applications. The miRNA circuits described here achieve precise, sequence-tunable gene dosage-invariant control of protein expression (**Figure 2.2 and 2.3**), orthogonal control of multiple target genes (**Figure 2.4**), and portability across cell types and modes of delivery (**Figure 2.5**). We therefore anticipate that they could become standard systems for controlled gene expression in diverse areas of biomedical science and biotechnology, including imaging (**Figure 2.6**), CRISPR-based gene editing (**Figure 2.7**), and AAV-based gene therapy (**Figure 2.7, Figure S2.7**).

The process of engineering these circuits revealed unexpected design principles.

Strong miRNA regulation, as obtained with the 1×21nt circuit, was not sufficient for dosage invariance within targeted expression levels (**Figure 2.1F**). Rather, simultaneously reducing complementarity and multimerizing target sites to engage multivalent TNRC6-dependent regulation was essential. In other regulatory systems, multispecific recognition is associated with ultrasensitivity. Here, however, it allowed linearly sensitive repression of mRNA required for dosage invariance but shifted responses to higher dosages. It will be interesting to learn whether natural miRNA regulatory systems use multispecific binding in similar ways [58, 59]. We also observed that separation of miRNA and target gene into divergently transcribed genes can be useful in allowing strong, independent control of miRNA expression relative to target mRNA (**Figure 2.3B-E**) but is not required for precise expression, facilitating applications like gene therapy where vector capacity can be limiting (**Data S2**).

The system may be extended or improved in different ways. First, better miRNA sequence-function models could potentially allow predictive design of setpoints. Second, current expression distributions exhibit significant variability, or noise (**Figure 2.3E**). This could reflect transcriptional bursting of the target or the miRNA, and could also be exacerbated by potential differences in the time delays for miRNA production and processing compared to RNA splicing and nuclear export. Going forward, understanding these and other contributions to overall variability could help to reveal fundamental limits of expression precision in the cell [4, 5, 60, 61]. Finally, the ability to combine DIMMER regulation with inducible promoters or natural enhancers for cell-type specificity could make these systems even more useful.

The circuits introduced here will be useful in diverse settings. They reduced background in imaging applications (**Figure 2.6**). They also reduced off-target RNA editing by CRISPR base editors, while maintaining on-target editing, suggesting they may be useful for gene editing applications (**Figure 2.7**). Analysis of stronger circuits with even lower expression setpoints could help to achieve even lower off-target edit rates. A major application category is gene therapy (**Figure 2.7G-H, Figure S2.7**). Many monogenic diseases that are gene therapy targets exhibit toxicity at high levels of the therapeutic gene, making it critical to suppress overexpression [62, 63, 64, 65, 66]. In the future, DIMMER circuits could also ensure fixed expression levels for receptors or other components in cell therapies [67], and therefore allow expression of transcription factors and other components at physio-

logical expression levels for regenerative medicine and other applications [68, 69]. Thus, we anticipate these systems becoming useful components in a wide range of engineered research and therapeutic contexts.

2.4.1 Limitations of the study

The DIMMER circuits have limitations. First, we do not yet have a predictive model of how an arbitrary miRNA sequence will behave quantitatively in a dosage compensation circuit. Varying miRNA regulation strength by modulating complementarity was straightforward with miR-L, but more complex for other synmiRs, suggesting that additional factors besides base pairing are likely important for fully predicting the activity of a miRNA on its target. Second, some aspects of the relevant molecular mechanisms remain unclear, including the role of catalytic slicer-dependent regulation [70, 71].

2.5 Supplemental Information

2.5.1 Supplemental Figures

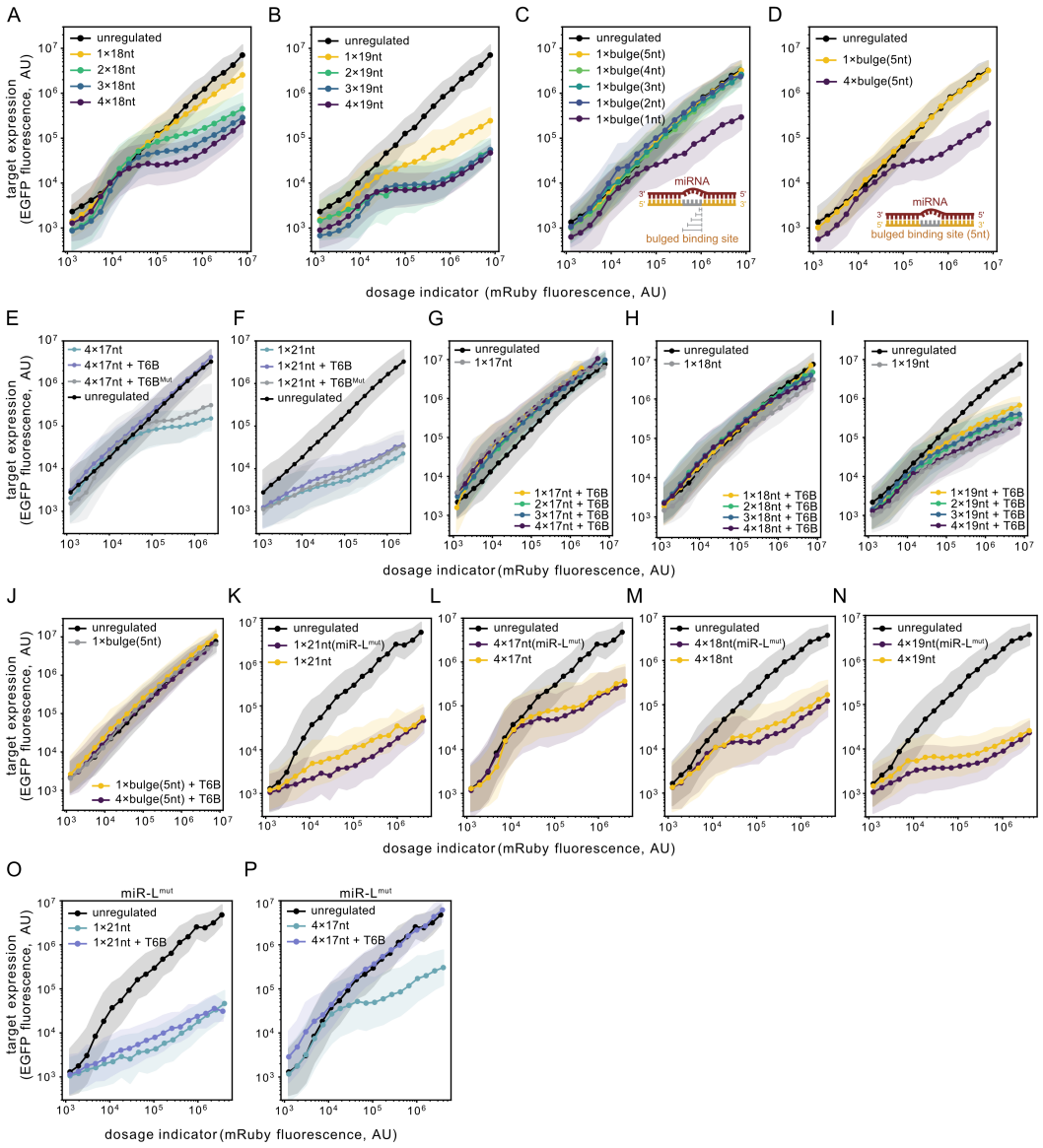

Figure S2.1: The dosage response curves of the miR-L and miR-L^{mut} -based targets and their dependencies of the TNRC6. Related to Figure 2.1 and Figure 2.2.

(A-B) The performance of the miR-L-mediated 1-4×18nt constructs **(A)** and 1-4×19nt constructs **(B)**, measured by flow cytometry.

(C-D) The performance of the bulged target constructs either in the single copy **(C)** or in the multiple copy **(D)** configurations. The bulge designs were shown as the insets in the center of the plots. The bulge region starts from the 9th nucleotide on the target, and the lengths of which range from 1 nt to 5 nt.

(E-F) We performed flow cytometry on the cells transfected with the miR-L 4×17nt construct **(E)** or miR-L 1×21nt construct **(F)** described in **Figure 2.2B-C** along with the fluorescent protein-only negative control, the T6B peptide, or the catalytically dead T6B peptide (denoted as T6B^{Mut}), respectively. The design of the catalytically dead T6B peptide was consistent with the previous literature [16].

Figure S2.1: (continued)
(G-J) We performed flow cytometry on the cells co-transfected with the labeled constructs and the T6B peptide or the control described in **Figure 2.2B-C** to identify the TNRC6 dependence of different constructs. T6B restored the multiple-site targets to the similar level of the corresponding single-site targets.
(K-N) We performed flow cytometry on the cells transfected with the constructs regulated by the miR-L^{mut}, and compared them with the corresponding constructs regulated by the original miR-L. The inclusion of the central mismatch in the miRNA slightly enhanced the regulation.
(O-P) We performed flow cytometry to measure the TNRC6 dependence of the 1×21nt **(O)** and the 4×17nt **(P)** constructs, respectively. Similar TNRC6 dependence was observed for these miR-L^{mut}-regulated constructs compared to that of the original miR-L-regulated constructs shown in **Figure 2.2B-C**.

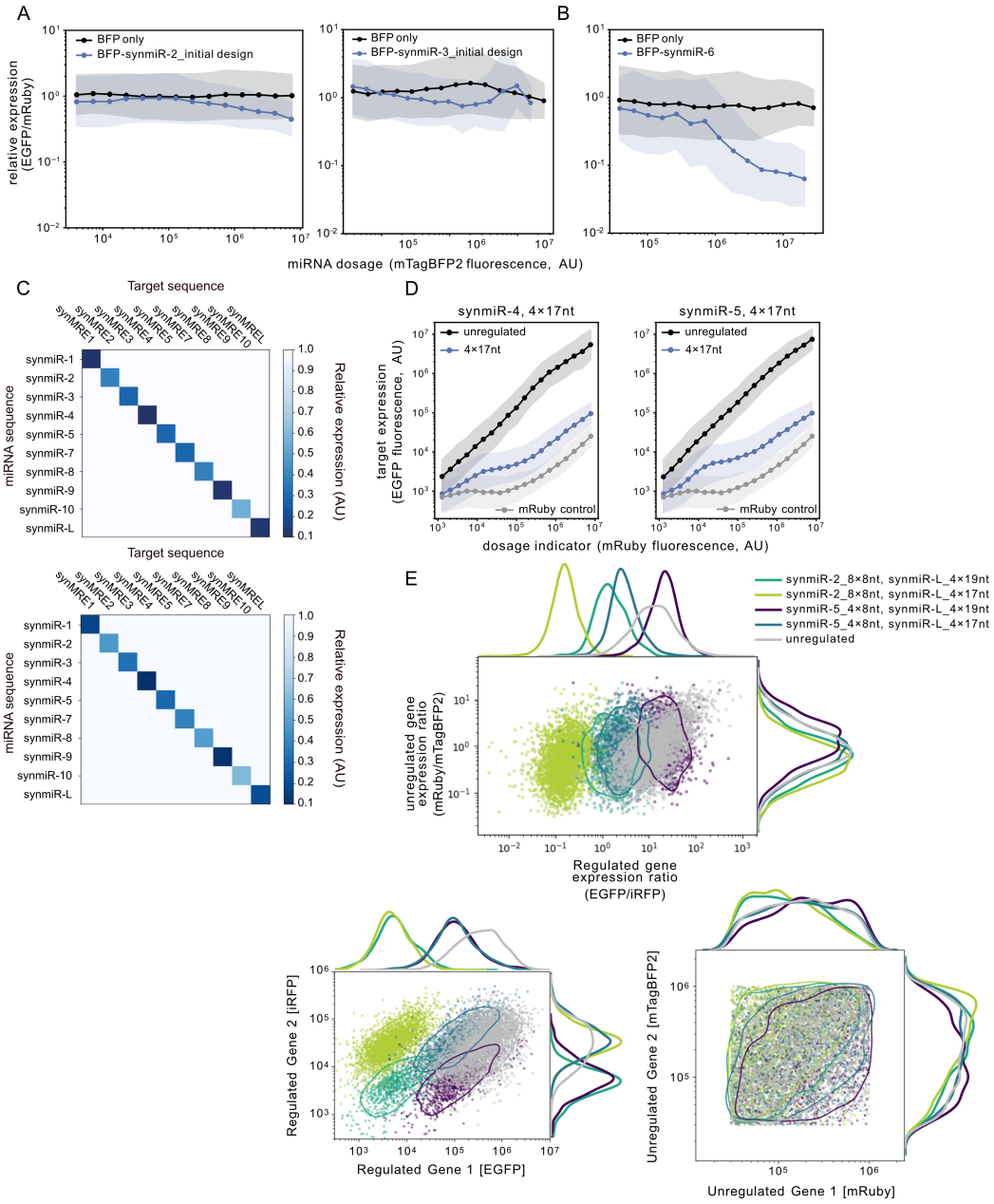

Figure S2.2: The initial synmiR-2, 3, and 6 performances, the biological replicates of the orthogonality test, the 4×17nt designs of synmiR-4 and synmiR-5, and the measurements of the fluorescent protein expression of the dual-reporter system. Related to Figure 2.4.

(A-B) We performed flow cytometry on the cells co-transfected with the circuit shown in **Figure 2.4A** using the initial sequence of synmiR-2 (**A, left panel**), synmiR-3 (**A, right panel**), synmiR-6 (**B**), and their corresponding targets with a single, fully complementary target site, respectively. The BFP only control does not have the 3'UTR miRNA. Relative expression levels were quantified as described in **Figure 2.4B**. Initial designs of synmiR-2 and synmiR-3 didn't function. synmiR-6 functioned but showed a sequence similarity to the endogenous miRNA hsa-mir-5697 (**STAR Methods**).

Figure S2.2: (continued)
(C) The other two biological replicates of the experiment described by **Figure 2.4E**.
(D) The dosage response curves of 4×17nt designs of synmiR-4 and synmiR-5, measured by flow cytometry.
(E) The biological replicate of the experiment described in **Figure 2.4H**, along with the unregulated control group (grey). The unregulated group contains the dual reporter system, both without the regulation element. Other experimental settings are the same as described in **Figure 2.4H**.
Upper panel, the distributions of the ratio of both unregulated proteins ([mRuby]/[mTagBFP2]) against the ratio of both regulated proteins ([EGFP]/[iRFP]).
The bottom left panel shows the regulated proteins' distributions of each group. The bottom right panel shows the dosage indicators' distributions of each group.
The dosage indicators exhibit the same distribution among different groups, by contrast, different combinations of DIMMERs allow a clear separation of the populations by the regulated proteins. Additionally, DIMMERs tightly control the stoichiometry of the regulated proteins.

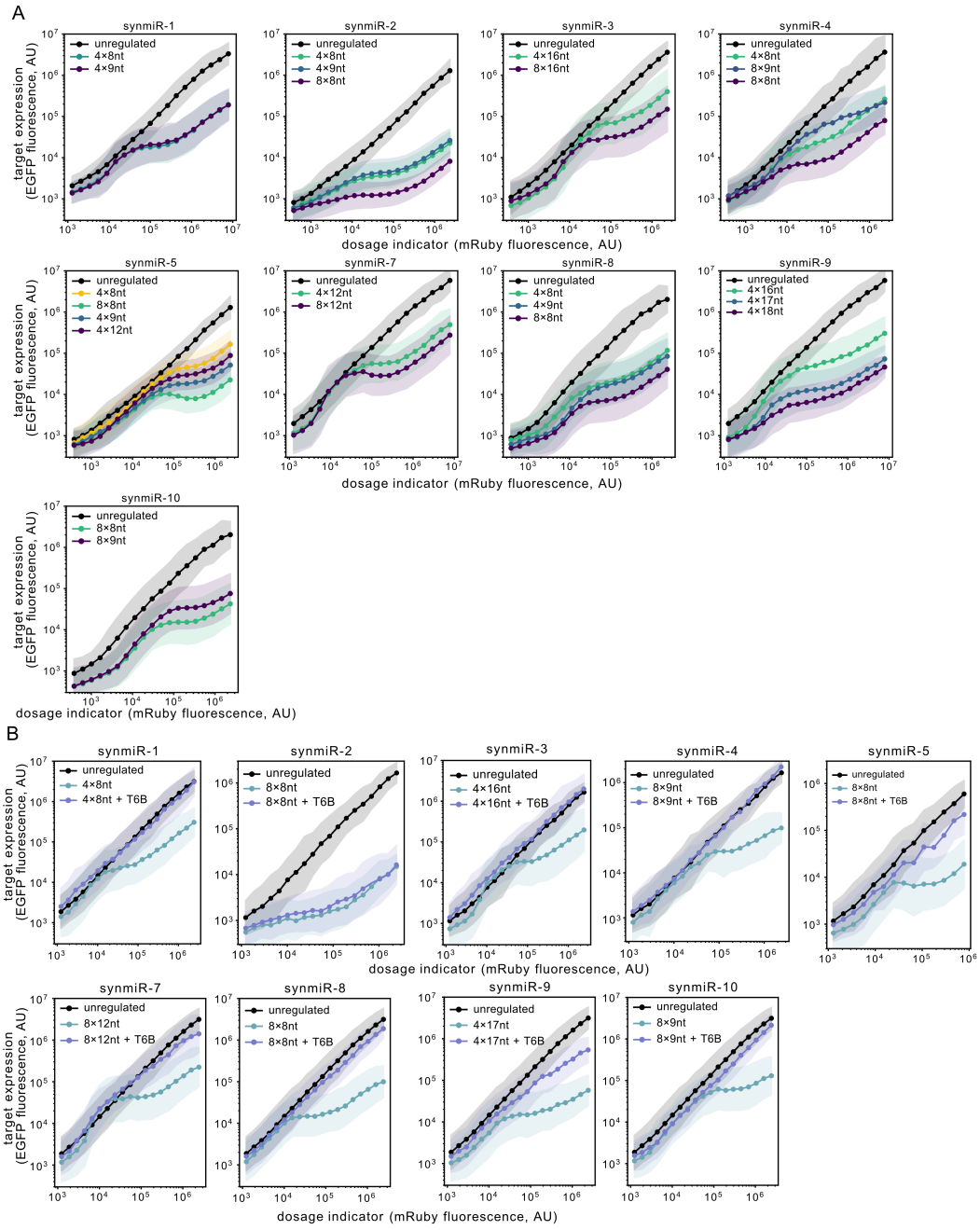

Figure S2.3: Figure S3. A gallery of all the DIMMER designs based on different miRNAs and targets (A) and the dependence of TNRC6-based regulation (B). Related to Figure 2.4.

Experimental settings are the same as described in **Figure 2.2B**. Almost all circuits used here relied on the TNRC6 to implement the inhibition, except synmiR-2 8×8nt, which might already be strong enough to achieve strong regulation.

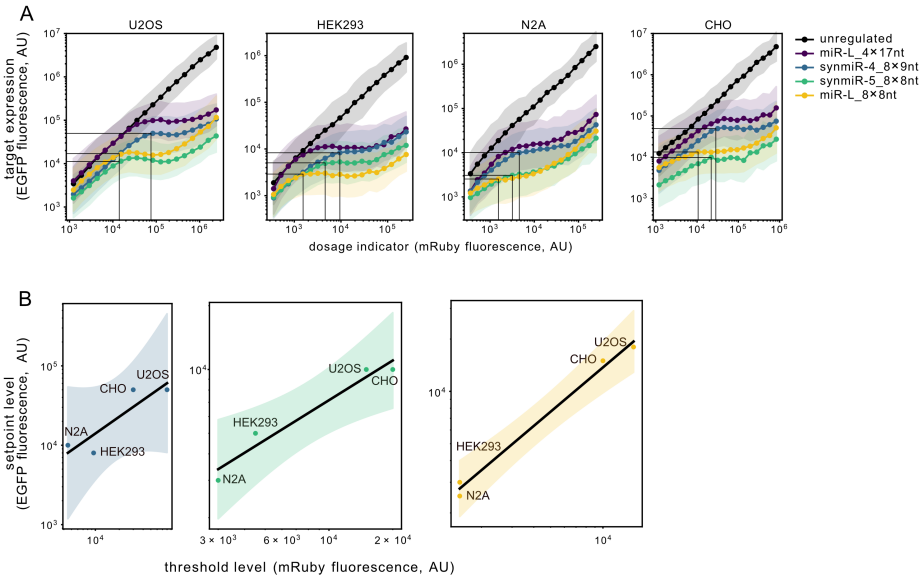

Figure S2.4: IFFL works across different cell lines. Related to Figure 2.5.
(A) The performance of different DIMMERs among various cell lines. The starting dosage of the dosage compensation behavior and the setpoint expression were indicated by the gray vertical lines and the horizontal lines, respectively.
(B) The threshold and the setpoint level co-vary across cell lines. The black solid line indicates a linear fit in the logarithm space (**STAR Methods**). See also **Figure 2.5B**.

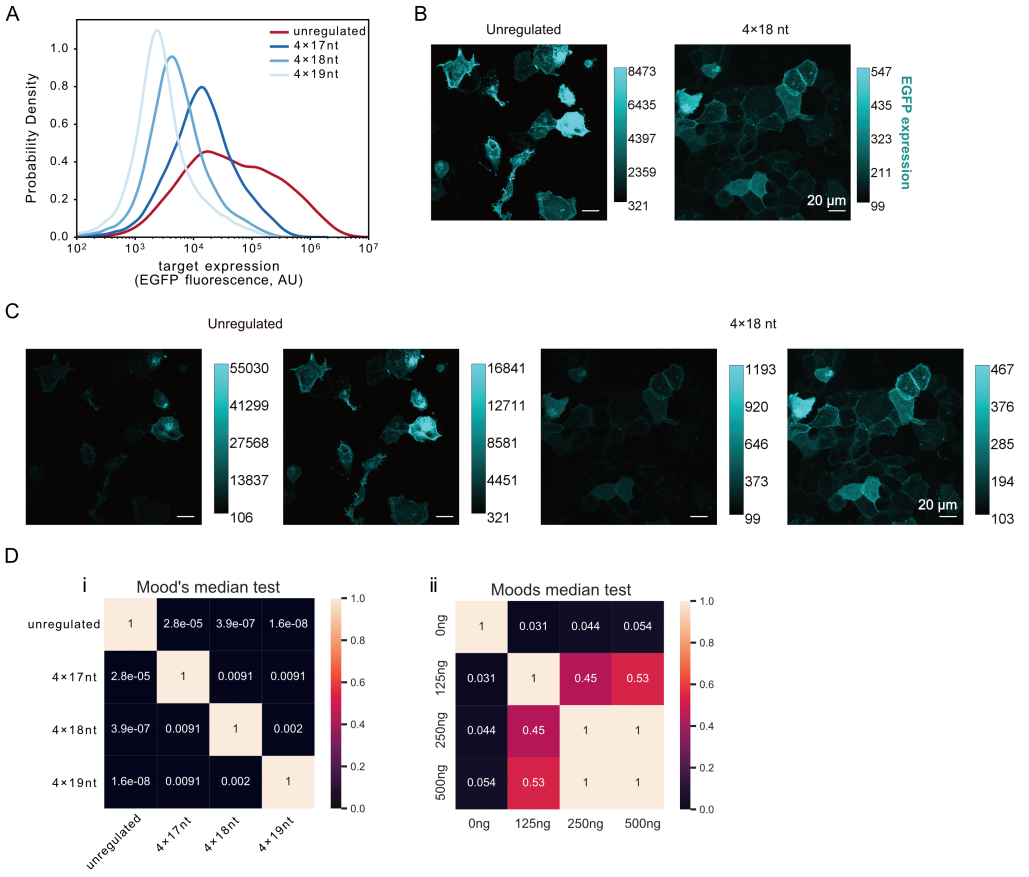

Figure S2.5: DIMMER circuit improves the DNA-PAINT experiment. Related to Figure 2.6.

(A) We performed flow cytometry on cells that were transfected with the EGFR-mEGFP with or without the DIMMER module. Cells were gated and binned by mRuby3 intensities. The expression of EGFR-mEGFP was plotted.

(B-C) Representative confocal microscopy images of the U2OS cells transfected with or without the DIMMER circuits **(B)** with different contrasts (**C**, indicated by the side color bars). The images were taken using the 60× magnification objective. Numbers on the color bars indicate the fluorescence intensities measured by imageJ. Scale bar, 20 μm. DIMMERs allow a more uniform expression at a lower setpoint.

(D) The statistical test of the DNA-PAINT experiment described in **Figure 2.6E** (i) and **Figure 2.6F** (ii). The pipeline of the statistical test is described in the **STAR Methods** section.

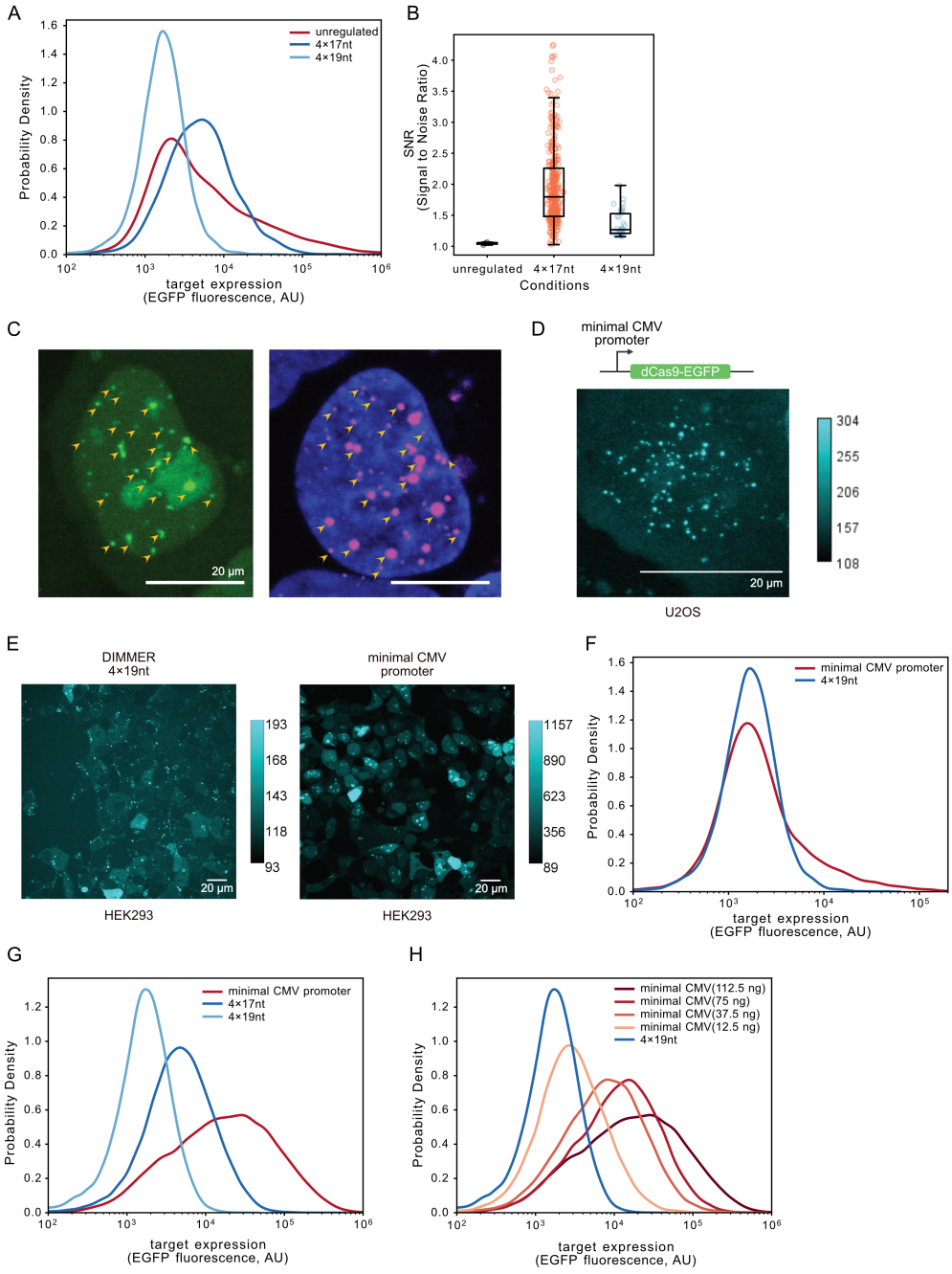

Figure S2.6: DIMMER circuit improves the CRISPR-based imaging. Related to Figure 2.6.

(A) We performed flow cytometry on the cells co-transfected with the dCas9-EGFP and the guide RNA plasmid, with or without the DIMMER circuit regulating the dCas9-EGFP. The plot shows the distributions of the EGFP fluorescence signal.

Figure S2.6: (continued)
(B) The boxplot shows the quantification of the signal-to-noise ratio (SNR) of the dots in the cells transfected with the dCas9-EGFP with or without the DIMMER module. Each scatter represents one dot inside a cell.
(C) The confocal microscopy images of one live U2OS cell expressing the dCas9-EGFP regulated by the 4×17nt circuit (left) and the same cell labeled by the DNA-FISH (right) probes targeting the telomeres post-fixation. The yellow arrowheads indicate the locations of the corresponding telomeres in the live cell and the fixed cell. Scale bar, 20 μm. Images were taken using the 60× objective. Detailed experimental procedures are described in **STAR Methods**.
(D) The confocal microscopy image of one live U2OS cell expressing the dCas9-EGFP driven by the minimal CMV promoter. Numbers on the color bars indicate the fluorescence intensities measured by imageJ. Scale bar, 20 μm.
(E) The confocal microscopy images of live HEK293T cells expressing the dCas9-EGFP either driven by the constitutive promoter and regulated by the 4×19nt DIMMER circuit (left) or driven by the minimal CMV promoter (right). Numbers on the color bars indicate the fluorescence intensities measured by imageJ. Scale bar, 20 μm.
(F) We performed flow cytometry on the U2OS cells expressing the dCas9-EGFP either driven by the constitutive promoter and regulated by the 4×19nt DIMMER circuit or driven by the minimal CMV promoter. The plot shows the distributions of the EGFP fluorescence signal.
(G) We performed flow cytometry on the HEK293T cells expressing the dCas9-EGFP either driven by the constitutive promoter and regulated by the 4×17nt and 4×19nt DIMMER circuits or driven by the minimal CMV promoter. The minimal CMV promoter generates a broader EGFP distribution.
(H) We performed flow cytometry on the HEK293T cells expressing the dCas9-EGFP either driven by the constitutive promoter and regulated by the 4×19nt DIMMER circuit or driven by the minimal CMV promoter with different transfection doses. The lowest transfection amount condition (12.5 ng) still cannot generate a distribution matching that of the 4×19nt DIMMER circuit.

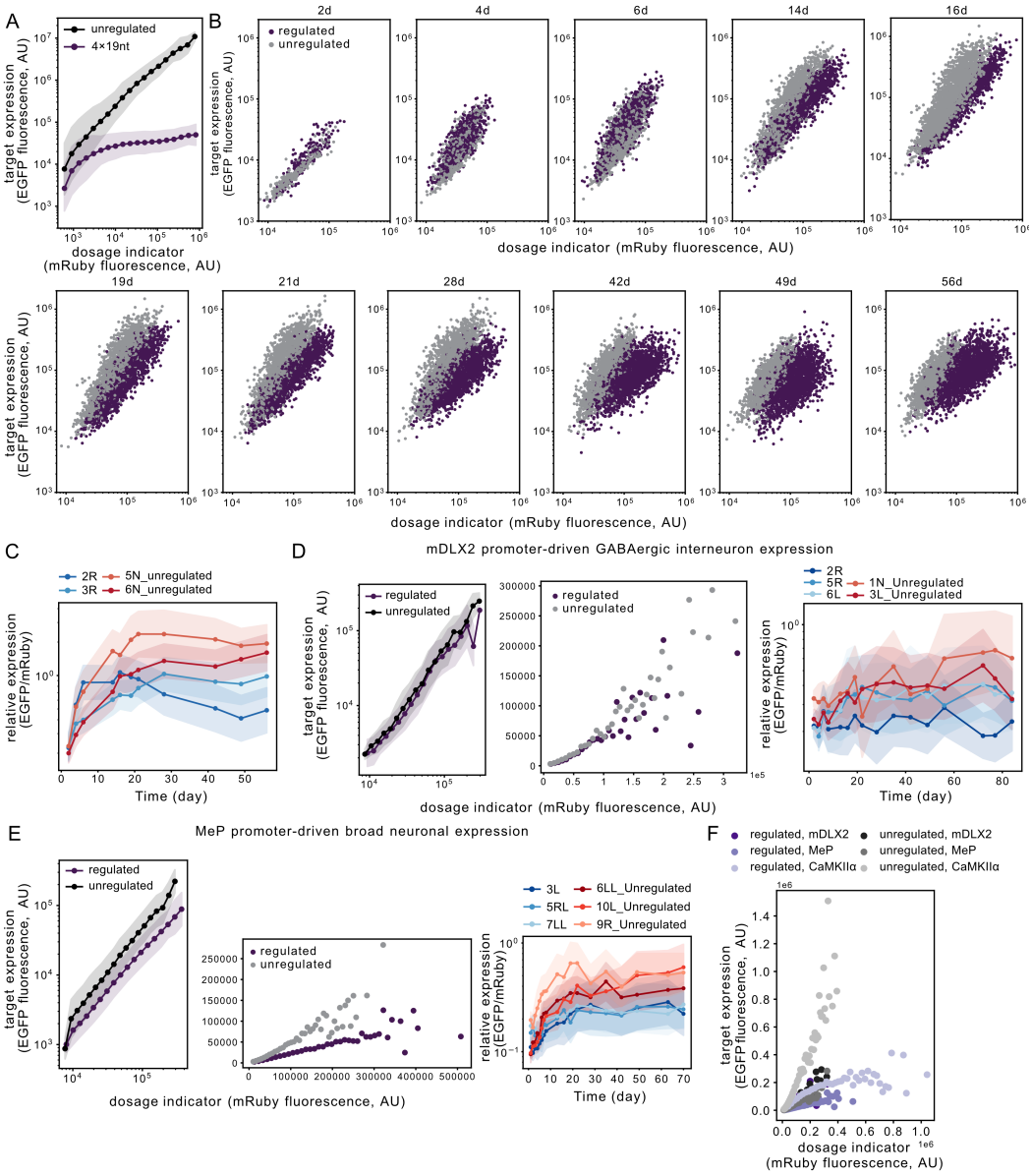

Figure S2.7: DIMMER circuits operate in live mouse brains. Related to Figure 2.7.

(A) We performed flow cytometry on the U2OS cells that were co-transfected with the circuits described in **Figure 2.7G**, upper panel.

(B) We quantified the expressions of H2B-mRuby3 and H2B-EGFP driven by the CaMKIIα promoter which labels the excitatory neurons over time. Each dot represents the protein expressions of one single cell. Expressions of the unregulated and the regulated groups both accumulated with time, and gradually showed separation from each other. The time shown on the top of each plot indicates the time post injection. The image analysis procedure is described in detail in **STAR Methods**.

Figure S2.7: (continued)
(C) We plotted the relative expression (the median single-cell EGFP/mRuby3 value) of each mouse at each single time point in the CaMKIIα promoter cohort. The orange curve and the red curve show the EGFP/mRuby dynamics of two mice in the unregulated group, and the light blue and the dark blue curves show the EGFP/mRuby dynamics of two mice in the 4×19nt DIMMER-regulated group. The shaded region was calculated as described in **Figure 2.1F**.
(D)(E) We used the mDLX2 promoter to drive the GABAergic interneuronal expression **(D)**, and used the MeP promoter to drive pan-neuronal expression **(E)** in the mouse brains. The left and the middle panels show the quantification of the pooled cells under each experimental condition, as described in **Figure 2.7H**, with the corresponding tissue-specific promoter, respectively. The right panels show the relative expression of each mouse at each single time point, as described in **(C)**, with the corresponding tissue-specific promoter, respectively.
(F) We pooled all the data of different cohorts together and plotted the EGFP intensities against mRuby3 intensities in the linear scale. The dots were calculated by the binned mRuby3 intensities.

2.5.2 Data S1: Mathematical modeling of the miRNA-based gene dosage compensation circuits, related to Figure 2.1.

I. Modeling miRNA-based gene dosage compensation circuits

Here we develop a simplified mathematical model of miRNA-based incoherent feed-forward loop (IFFL) circuits, and use it to explore how various parameters control the scaling of target gene expression with gene dosage. The focus in this section is to understand how general properties of the circuit influence its ability to perform dosage compensation. We also include a second Supplementary modeling section that focuses on the role of specific molecular mechanisms.

In this model, we focused on three sets of dynamic processes: RISC production and removal, mRNA production and removal, and RISC-mRNA complex formation.

1. RISC production and removal. We assume miRNA is expressed, processed, and loaded into Argonaute (Ago) proteins to form an active (mature) RNA-Induced Silencing Complex (RISC) [72]. We denote the concentration of mature RISC (containing the miRNA) as r. We assume RISC is produced at a total rate of $D\beta_{\mathrm{r}}$, where D denotes gene copy number (gene dosage) and β_{r} denotes the rate of production RISC production per gene copy. This expression implicitly assumes that miRNA expression levels do not saturate available miRNA processing machinery, Ago, or other components. We also assume that the RISC complex is removed at

total rate $\gamma_r r$, where γ_r denotes a combined rate constant for dilution, degradation, and other removal processes.

2. mRNA production and removal. We assume mRNA, denoted m, is produced at a rate proportional to gene copy number, $D\beta_m$, where β_m is the mRNA production rate per gene copy. mRNA can also be removed at a rate $\gamma_m m$ due to dilution and degradation. We note that even though mRNA and miRNA are produced from the same engineered locus, the production rate constants β_m and β_r can differ, since the miRNA and mRNA are produced from distinct promoters, and/or are processed through different downstream pathways.

3. RISC-mRNA complex formation and dissociation. We assume that the RISC and target mRNA associate to form a complex, whose concentration is denoted C, at a rate $k_{on} r m$, following mass action kinetics with rate constant, k_{on}. Once formed, this complex can dissociate at rate $k_{off} C$, undergo catalytic mRNA degradation, at rate $k_c C$.

These chemical reactions can be summarized as follows, using G to denote the gene, present at copy number (dosage) D:

$$
\begin{aligned}
G &\xrightarrow{D\beta_r} G + r \\
r &\xrightarrow{\gamma_r} \emptyset \\
G &\xrightarrow{D\beta_m} G + m \\
m &\xrightarrow{\gamma_m} \emptyset \\
r + m &\xrightarrow{k_{on}} C \\
C &\xrightarrow{k_{off}} r + m \\
C &\xrightarrow{k_c} r
\end{aligned}
$$

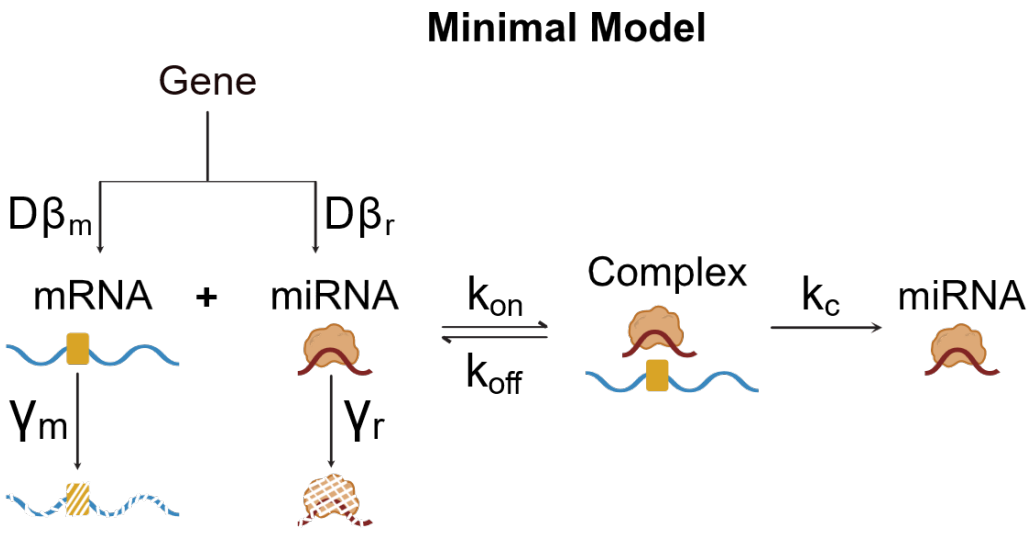

Figure S2.8: Reactions in the minimal model, including mRNA and miRNA production and removal, as well as RISC complex formation and catalytic mRNA degradation. Parameters on arrows denote reaction rates.

With these definitions and assumptions, we can write down a set of ordinary differential equations for the three variables, r, m, and C.

$$\frac{dr}{dt} = \beta_r D - \gamma_r r - k_{on} r m + (k_c + k_{off}) C$$
$$\frac{dm}{dt} = \beta_m D - \gamma_m m - k_{on} r m + k_{off} C$$
$$\frac{dC}{dt} = k_{on} r m - (k_c + k_{off}) C$$

These equations resemble previous modeling of regulation by small RNAs [73, 74], except for the explicit incorporation of gene dosage in the control of both mRNA and miRNA.

To gain insight into the possible behaviors of this system, we first non-dimensionalize it. We define a dimensionless time, $\widetilde{t} = \gamma_r t$, by rescaling time in units of the RISC lifetime. We also define a dimensionless RISC concentration, $\widetilde{r} = r/(\beta_r/\gamma_r)$. This effectively rescales r in units of the unregulated steady-state expression level produced by a single copy of the gene. We similarly define a dimensionless mRNA concentration, $\widetilde{m} = m/(\beta_m/\gamma_m)$, normalizing m by its single copy steady state expression level. Finally, we define a dimensionless concentration of the RISC-miRNA-mRNA complex, $\widetilde{C} = C/(\beta_m/\gamma_m)$.

In addition, we define a set of convenient dimensionless parameter ratios:

$$\gamma = \frac{\gamma_{\mathrm{m}}}{\gamma_{\mathrm{r}}}$$
$$\beta = \frac{\beta_{\mathrm{m}}}{\beta_{\mathrm{r}}}$$
$$\widetilde{k}_{\mathrm{on}} = \frac{k_{\mathrm{on}}\beta_{\mathrm{r}}}{\gamma_{\mathrm{m}}\gamma_{\mathrm{r}}}$$
$$\widetilde{k}_{\mathrm{off}} = \frac{k_{\mathrm{off}}}{\gamma_{\mathrm{m}}}$$
$$\widetilde{k}_{\mathrm{c}} = \frac{k_{\mathrm{c}}}{\gamma_{\mathrm{m}}}$$
$$K = \frac{\widetilde{k}_{\mathrm{c}} + \widetilde{k}_{\mathrm{off}}}{\widetilde{k}_{\mathrm{on}}}$$

In the non-dimensionalized system, the differential equations can be written as,

$$\frac{d\widetilde{r}}{d\widetilde{t}} = D - \widetilde{r} - \widetilde{k}_{\mathrm{on}}\beta\widetilde{m}\widetilde{r} + (\widetilde{k}_{\mathrm{c}} + \widetilde{k}_{\mathrm{off}})\beta\widetilde{C}$$
$$\gamma^{-1}\frac{d\widetilde{m}}{d\widetilde{t}} = D - \widetilde{m} - \widetilde{k}_{\mathrm{on}}\widetilde{m}\widetilde{r} + \widetilde{k}_{\mathrm{off}}\widetilde{C}$$
$$\gamma^{-1}\frac{d\widetilde{C}}{d\widetilde{t}} = \widetilde{k}_{\mathrm{on}}(\widetilde{m}\widetilde{r} - K\widetilde{C})$$

By definition, at steady state, the time derivatives all equal zero. Denoting steady values with a subscript s, we then have:

$$D - \widetilde{r}_{\mathrm{s}} - \widetilde{k}_{\mathrm{on}}\beta\widetilde{m}_{\mathrm{s}}\widetilde{r}_{\mathrm{s}} + (\widetilde{k}_{\mathrm{c}} + \widetilde{k}_{\mathrm{off}})\beta\widetilde{C}_{\mathrm{s}} = 0$$
$$D - \widetilde{m}_{\mathrm{s}} - \widetilde{k}_{\mathrm{on}}\widetilde{m}_{\mathrm{s}}\widetilde{r}_{\mathrm{s}} + \widetilde{k}_{\mathrm{off}}\widetilde{C}_{\mathrm{s}} = 0$$
$$\widetilde{m}_{\mathrm{s}}\widetilde{r}_{\mathrm{s}} - K\widetilde{C}_{\mathrm{s}} = 0$$

Solving the equations above, we obtain an equation for steady-state mRNA concentration:

$$\widetilde{m}_{\mathrm{s}} = \frac{D}{1 + \frac{\widetilde{k}_{\mathrm{c}}}{K}D}$$

Henceforth, we will omit the tildes for notational convenience, and switch to the non-dimensionalized variables and parameters.

In the limit of large dosage, $\frac{k_{\mathrm{c}}}{K}D \gg 1$. As a result, m_{s} approaches a limiting value, $m_{\mathrm{s}} \to \frac{K}{k_{\mathrm{c}}}$ independent of gene dosage. This is the regime that the constructs developed here are targeting.

Figure 2.1D plots this expression, and Figure S2.9A and B plot the expression in different paramters, i.e. k_{off}, and k_{c}.

Finally, to estimate biological parameter values, we assume that the miRNA and mRNA are produced at similar rates, using estimated parameter values derived from [75] with tunable ranges, as listed in the following table:

Non-dimensional parameter	Dimensionless values
β	1
γ	0.8
k_{on}	200000
k_{off}	10, 100, 1000, 10000
k_{c}	400, 40, 4, 0.4, 0.004, 0.0004

Table 2.1: Dimensionless parameters for the minimal model.

This simple model suggests that steady state mRNA concentration can—under some regimes—achieve dosage independence.

I.1 Incorporating ultrasensitivity

Multiple mechanisms can give rise to ultrasensitivity in miRNA regulation, which is not present in the model so far. To allow for ultrasensitivity, we considered a more general, phenomenological model in which target mRNA inhibition by miRNA follows a Hill function of the RISC concentration, with Hill coefficient n. This treatment omits intermediate steps, i.e., the RISC-mRNA complex formation and dissociation. With this assumption, we can write down a different set of ordinary differential equations for the two variables, r, and m.

$$\frac{dr}{dt} = \beta_{\text{r}} D - \gamma_{\text{r}} r$$
$$\frac{dm}{dt} = \frac{\beta_{\text{m}} D}{1 + (\frac{r}{\kappa})^n} - \gamma_{\text{m}} m$$

To non-dimensionalize this system, we define $\widetilde{t} = \gamma_{\text{r}} t$, $\widetilde{r} = r/(\beta_{\text{r}}/\gamma_{\text{r}})$, and $\widetilde{m} = m/(\beta_{\text{m}}/\gamma_{\text{m}})$. Additionally, we define $\widetilde{K} = \frac{\beta_{\text{r}}}{\gamma_{\text{r}}\kappa}$ for convenience. The steady state expression of mRNA concentration can then be written as:

$$\widetilde{m_{\text{s}}} = \frac{D}{1 + (\widetilde{K} D)^n}$$

Figure S2.9C plots this expression when $\widetilde{K} = 1$ and $n = 0.5, 1, 2$. Critically, only when $n = 1$, does $\widetilde{m}_{\text{s}}$ approach a limiting value independent of D. If $n < 1$, $\widetilde{m}_{\text{s}}$

shows a sublinear increase with D. If $n > 1$, $\widetilde{m}_{\mathrm{s}}$ shows a biphasic dependence on D (Figure S2.9C).

This analysis suggests that it is important to maintain a Hill coefficient of $n \sim 1$ for optimal dosage compensation.

I.2 Incorporating complex degradation and the bounded mRNA Thus far, we ignored possible degradation of mRNA and miRNA within the complex, and did not explicitly account for total mRNA, denoted m_{tot}, which includes both free mRNA and mRNA engaged in RISCs. Here, we incorporate these additional reactions and analyze their effects on the dosage response behavior. Specifically, we added two additional reactions:

$$C \xrightarrow{\gamma_r} m$$
$$C \xrightarrow{\gamma_m} r$$

We now write down the modified ordinary differential equations:

$$\begin{aligned}
\frac{dr}{dt} &= \beta_{\mathrm{r}} D - \gamma_{\mathrm{r}} r - k_{\mathrm{on}} r m + (k_{\mathrm{c}} + k_{\mathrm{off}} + \gamma_{\mathrm{m}}) C \\
\frac{dm}{dt} &= \beta_{\mathrm{m}} D - \gamma_{\mathrm{m}} m - k_{\mathrm{on}} r m + (k_{\mathrm{off}} + \gamma_{\mathrm{r}}) C \\
\frac{dC}{dt} &= k_{\mathrm{on}} r m - (k_{\mathrm{c}} + k_{\mathrm{off}} + \gamma_{\mathrm{m}} + \gamma_{\mathrm{r}}) C
\end{aligned}$$

We then perform non-dimensionalization, as above. The only dimensionless parameter that has been changed is K, which now is $K = \frac{\widetilde{k}_{\mathrm{c}} + \widetilde{k}_{\mathrm{off}} + 1 + \gamma^{-1}}{\widetilde{k}_{\mathrm{on}}}$, leading to the following dimensionless differential equations:

$$\begin{aligned}
\frac{d\widetilde{r}}{d\widetilde{t}} &= D - \widetilde{r} - \widetilde{k}_{\mathrm{on}} \beta \widetilde{m} \widetilde{r} + (\widetilde{k}_{\mathrm{c}} + \widetilde{k}_{\mathrm{off}} + 1) \beta \widetilde{C} \\
\gamma^{-1} \frac{d\widetilde{m}}{d\widetilde{t}} &= D - \widetilde{m} - \widetilde{k}_{\mathrm{on}} \widetilde{m} \widetilde{r} + (\widetilde{k}_{\mathrm{off}} + \gamma^{-1}) \widetilde{C} \\
\gamma^{-1} \frac{d\widetilde{C}}{d\widetilde{t}} &= \widetilde{k}_{\mathrm{on}} (\widetilde{m} \widetilde{r} - K \widetilde{C})
\end{aligned}$$

From this, we can compute steady state concentrations:

$$\widetilde{C_s} = \frac{\widetilde{m_s}\widetilde{r_s}}{K}$$

$$\widetilde{m_s} = \frac{D}{1 + \frac{(1+\widetilde{k_c})\widetilde{r_s}}{K}}$$

$$\widetilde{r_s} = \frac{D(\widetilde{k_c}+1) - \beta D\gamma^{-1} - K + \sqrt{(D(\widetilde{k_c}+1) - \beta D\gamma^{-1} - K)^2 + 4DK(\widetilde{k_c}+1)}}{2(\widetilde{k_c}+1)}$$

Considering the parameter regime where $\widetilde{k_c} + 1 \gg \beta\gamma^{-1}$, the $\beta D\gamma^{-1}$ term is negligible. Therefore, we get $\widetilde{r_s} \approx D$. Henceforth, we drop the tildes for notational convenience, as in the previous sections. We write down the expression for m_{tot}:

$$m_{tot} = m_s + C_s = m_s(1 + \frac{r_s}{K}) = \frac{D(K + r_s)}{K + (k_c + 1)r_s} \approx \frac{D(K + D)}{K + (k_c + 1)D}$$

We now discuss the behavior of m_{tot} at different dosages:

1. When $D \ll \frac{K}{k_c+1} < K$: $m_{tot} \to D$. In this case, the total mRNA scales linearly with the dosage, similar to the small dosage regime in the minimal model.

2. When $\frac{K}{k_c+1} < D \ll K$: $m_{tot} \approx \frac{D}{1+\frac{(k_c+1)D}{K}} \to \frac{K}{k_c+1}$. In this case, m_{tot} shows dosage compensation.

3. When $D \gg K$: $m_{tot} \to \frac{D}{k_c+1}$. In this case, the total mRNA is again approximately proportional to dosage, but has a lower setpoint compared with the unregulated condition due to repression. Figure S2.9D depicts this behavior.

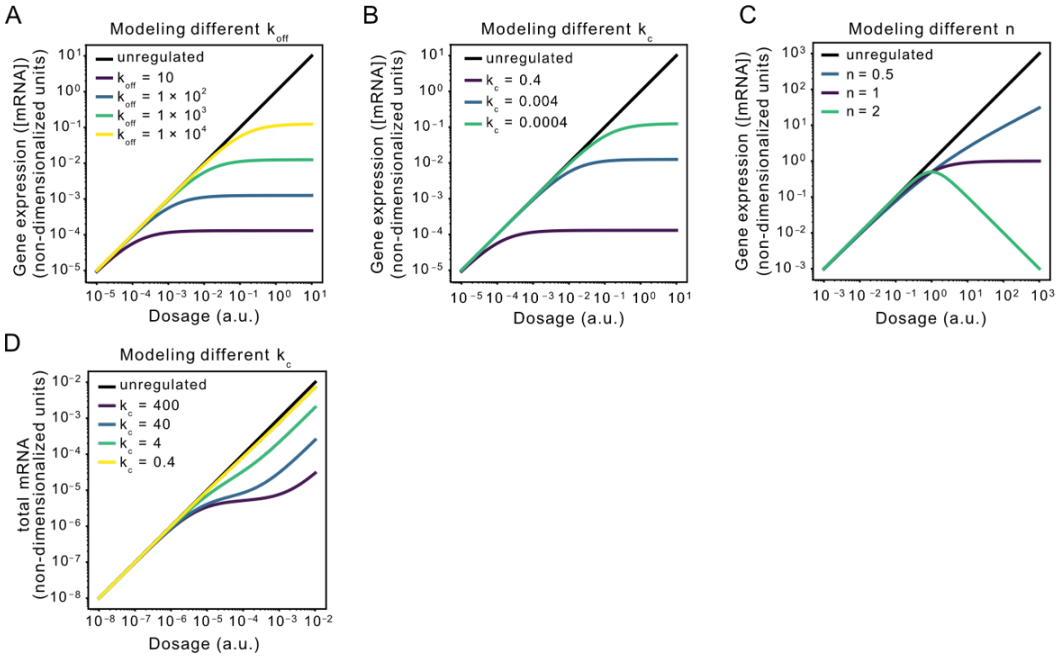

Figure S2.9: The minimal model of the miRNA-mediated IFFL. A, modeling the steady-state, dimensionless mRNA concentration under different k_{off} while maintaining $k_{\text{on}} = 2\times10^5$, $k_{\text{c}} = 0.4$. B, modeling the steady-state, dimensionless mRNA concentration under different k_{c} while maintaining $k_{\text{on}} = 2\times10^5$, $k_{\text{off}} = 10$. C, modeling the steady-state, dimensionless mRNA concentration under different Hill coefficient n while maintaining $K = 1$. D, modeling the steady-state, dimensionless total mRNA amount (including the free mRNA and the bounded complex) under different k_{c} while maintaining $k_{\text{on}} = 2\times10^5$, $k_{\text{off}} = 10$ by considering the complex degradation caused by natural degradation of either mRNA or miRNA, and the translational contribution of the bounded complex.

II. Conclusions from simple modeling of miRNA-based gene dosage compensation circuits

Based on the results above, we obtain the following main design guidelines for miRNA-based dosage compensation circuits:

1. miRNA-based circuits incorporating miRNA-mRNA complex formation and catalytic mRNA degradation should be able to achieve dosage compensation.

2. However, a critical requirement for dosage compensation is non-ultrasensitive regulation of mRNA by miRNA, i.e. a Hill coefficient $n \sim 1$.

3. Degradation of mRNA and miRNA within the RISC can affect dosage compensation. However, these reactions still allow dosage compensation within a limited regime provided that k_c is sufficiently high (Figure S2.9D). The cutoff for this be-

havior occurs when dosage levels increase enough to make r_s becomes comparable to K, leading to a "tail" in the expression versus dosage plot.

2.5.3 Data S2: Iterative engineering, auxiliary analyses of variance, alternative visualizations of the main figure datasets, and other technical controls, related to Figure 2.2, 2.3, 2.4, 2.5, 2.7.

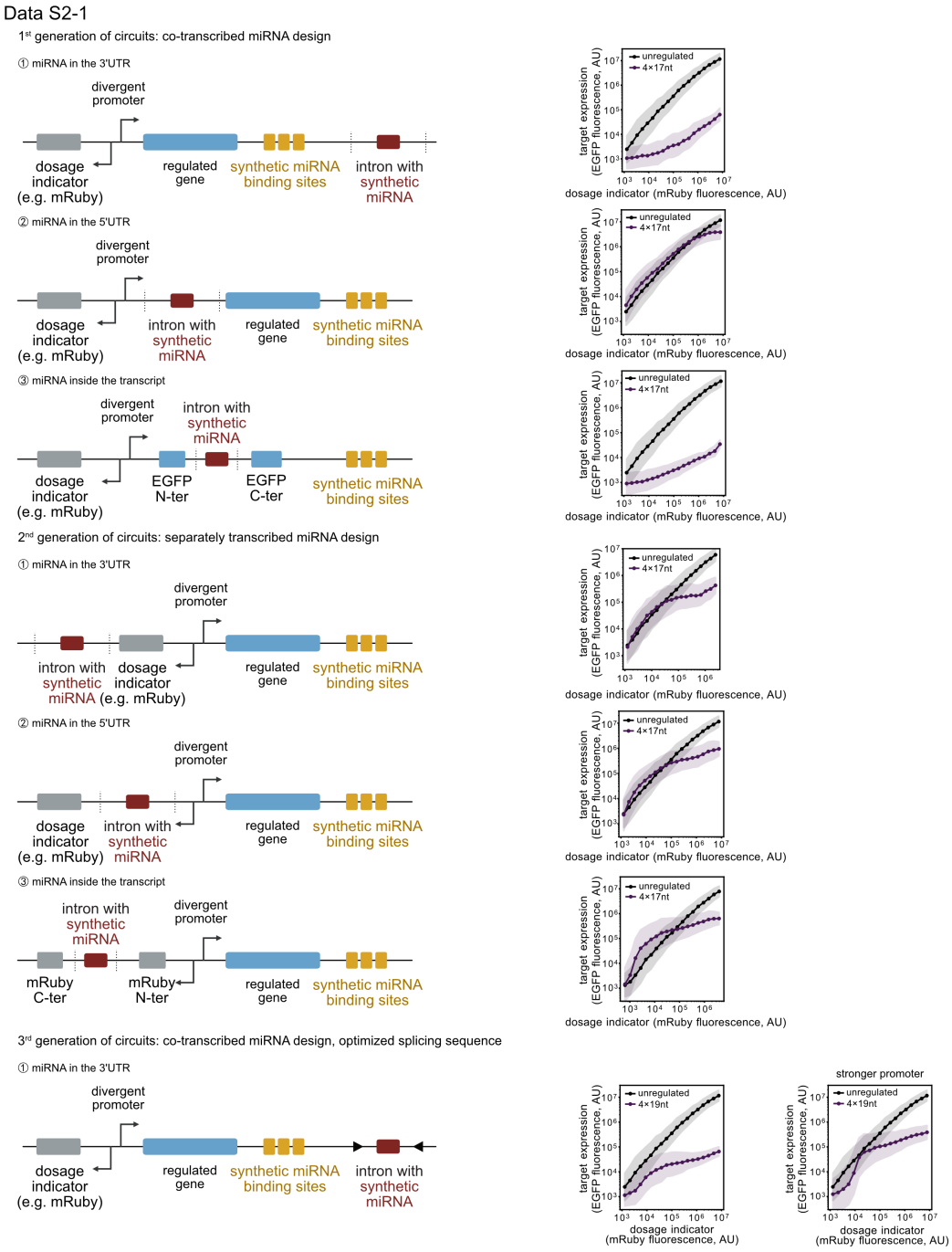

Figure S2.10: The iterative engineering process of different circuit architectures, related to Figure 2.2.

Left panel, various circuit configurations; right panel, corresponding circuit quantitative behaviors measured by flow cytometry.

Figure S2.10: (continued)
In the first generation of the circuit, a co-transcribed miRNA design was used. Placing the intronic synthetic miRNA in the 3'-UTR or in the middle of the transcript both lead to the strong repression effects that might be caused by the destabilization of the transcript that is independent of the miRNA inhibition. Placing the intronic synthetic miRNA in the 5'-UTR of the transcript does not produce enough miRNA. In the second generation of the circuit, the intronic synthetic miRNA sequence was placed in the separate transcript (which is mRuby3). No matter where the miRNA sequence is inserted, the dosage compensation behavior is maintained. However, when the miRNA sequence is placed into the 3'-UTR, the amount of the mature miRNA seems to be the greatest, since the circuit has the lowest setpoint.
In the third generation of the circuit, the intronic synthetic miRNA sequence was placed in the same transcript as the regulated EGFP. The key difference between the third generation design and the first generation design is that the splicing signal sequence is optimized (indicated by the black triangles). The dosage compensation behavior is achieved, and the utilization of a stronger promoter changes the setpoint.

Circuit configuration	**miRNA production**	**Transcript stability**
1st generation, ①	normal-to-low miRNA production	unstable
1st generation, ②	low miRNA production	stable
1st generation, ③	almost no miRNA production	unstable
2nd generation, ①	normal miRNA production	stable
2nd generation, ②	lower miRNA production compared to 2nd generation, ①	stable
2nd generation, ③	lower miRNA production compared to 2nd generation, ①	stable
3rd generation, ①	normal miRNA production	stable
3rd generation, ①, stronger promoter	normal miRNA production	stable

Table 2.2: Summary of performances of different circuit configurations in Data S2-1.

Data S2-2

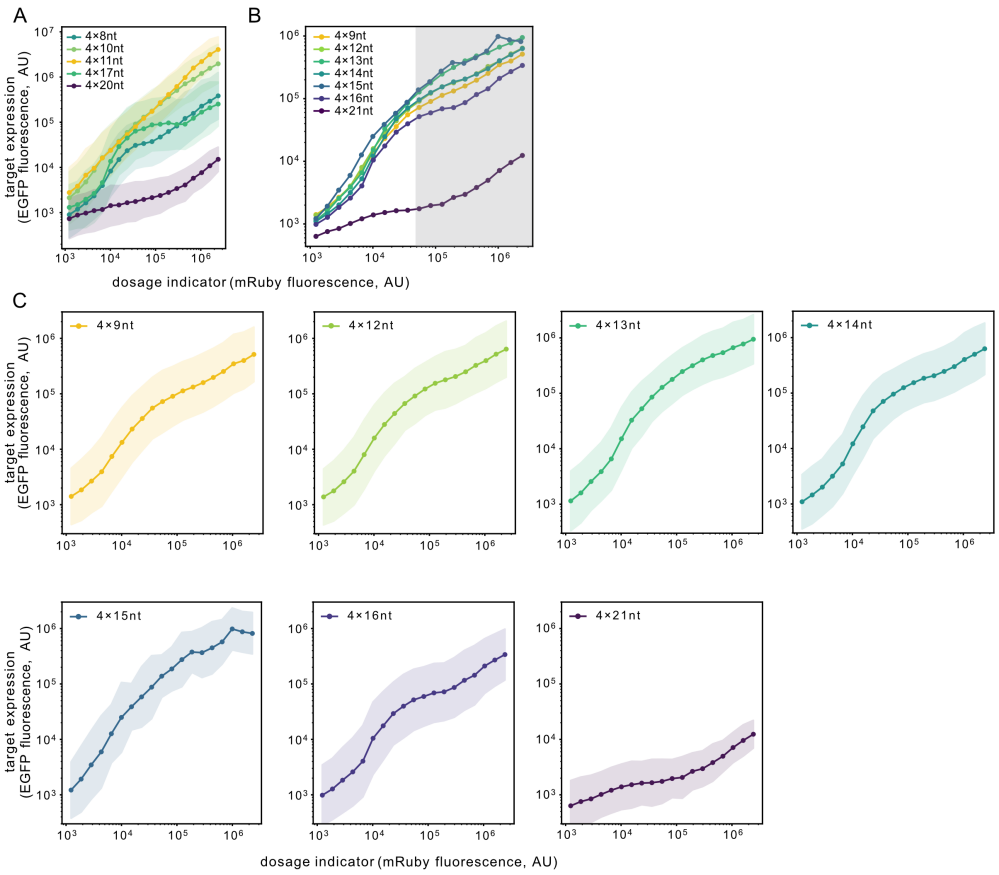

Figure S2.11: The dosage response curves of multiple miR-L 4×n circuits. Related to Figure 2.3.

(A) The curves in **Figure 2.3A** right panels with the geometric variances plotted.

(B) We performed flow cytometry on the cells transfected with the miR-L regulating 4×n circuits described in **Figure 2.3A**. Gray rectangle indicates the gated region of the **Figure 2.3A** left panel.

(C) The curves in **(B)** with the geometric variances were plotted separately to avoid being cluttered together.

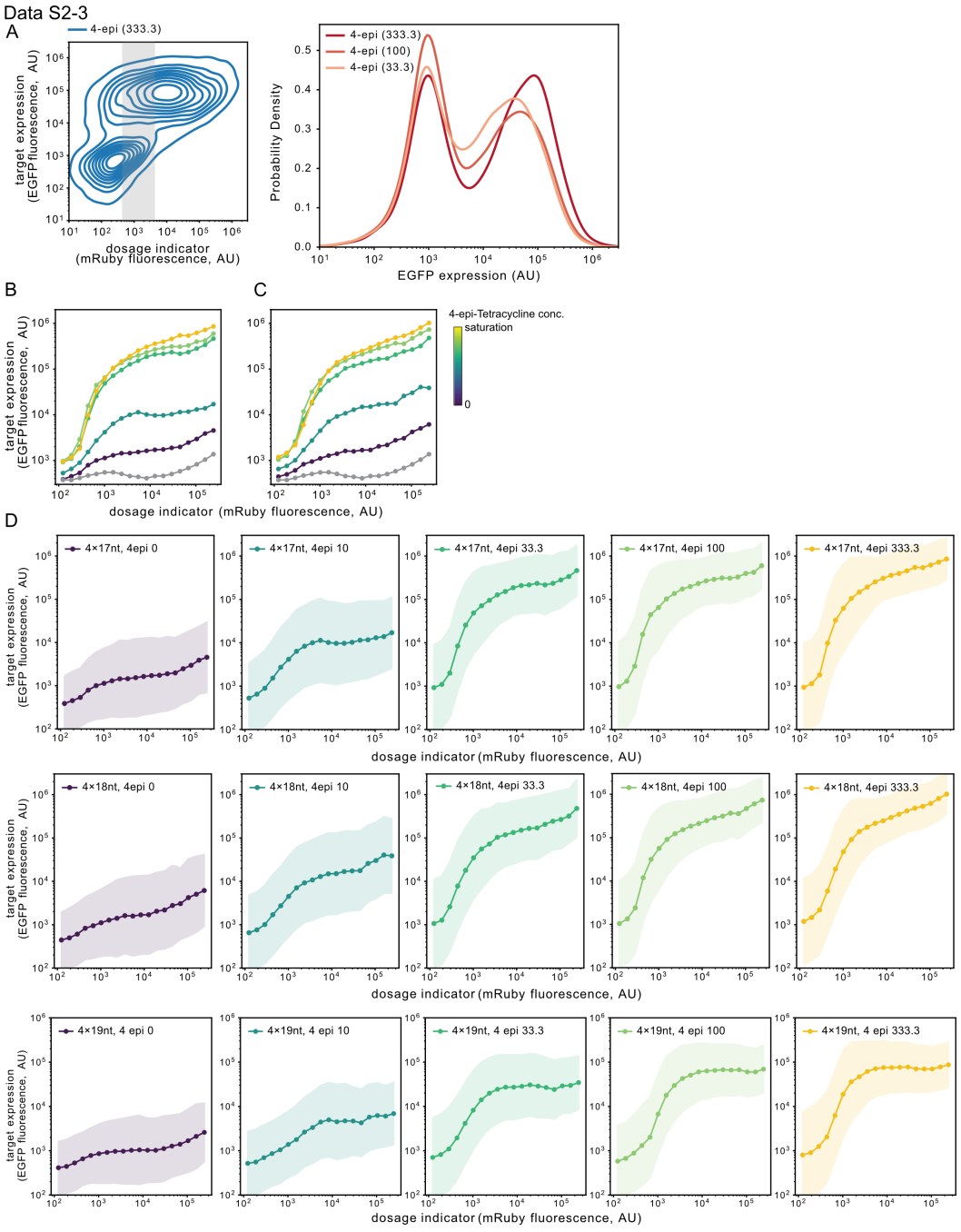

Figure S2.12: The dosage response curves of the inducible DIMMER circuits in TRex cells. Related to Figure 2.3.

(A) The left panel shows the contour plot of the 4×19nt construct described in **Figure 2.3E**. The concentration of 4-epi was 333.3 ng/mL. Cells falling into the shaded grey region were selected to plot the EGFP distribution in the right panel. Cells in **Figure 2.3E** falling into the same mRuby3 range were also selected to plot the EGFP distribution at the 4-epi concentration of 100 and 33.3 ng/mL. The bimodal distribution of EGFP indicates the bursty nature of the Tet-On promoter used here, and explains the seemingly ultrasensitive behavior observed in **Figure 2.3E** when the dosage is low.

Figure S2.12: (continued)
(B-C) We performed flow cytometry on the TRex cell line which was transfected with the 4×17nt **(B)** or 4×18nt **(C)** DIMMER construct described in **Figure 2.3E**. The concentrations of the 4-epi-Tetracycline, from purple to yellow, were 0, 10, 33.3, 100, 333.3 ng/mL. The gray curve denotes the mRuby-only transfection control.
(D) The curves in **Figure 2.3E** and **Data S2-3 B-C** with the geometric variances plotted separately to avoid being cluttered together.

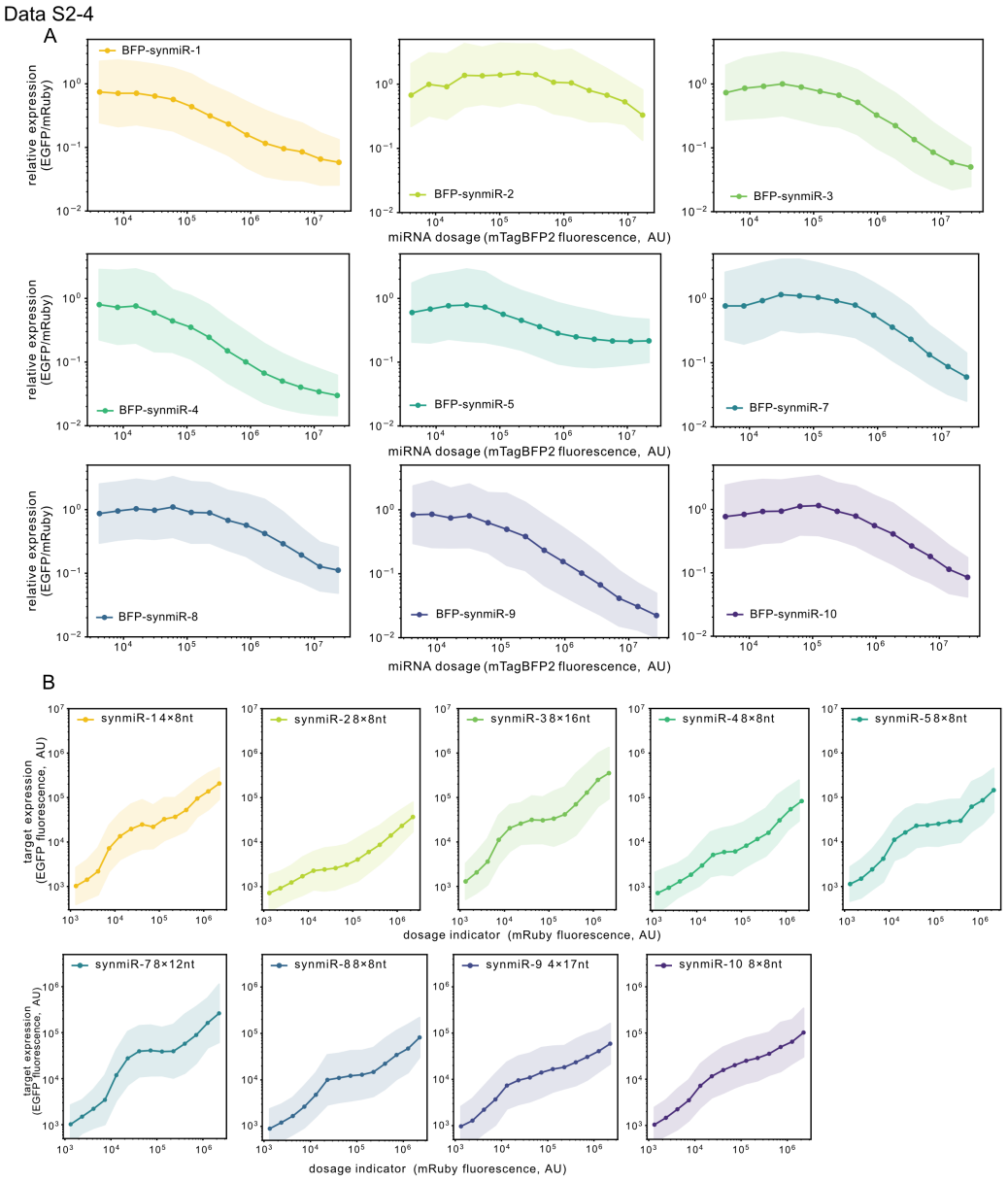

Figure S2.13: The geometric variance of the dosage response curves in Figure 2.4D (A) and Figure 2.4F (B) to avoid curve cluttering. Related to Figure 2.4.

Data S2-5

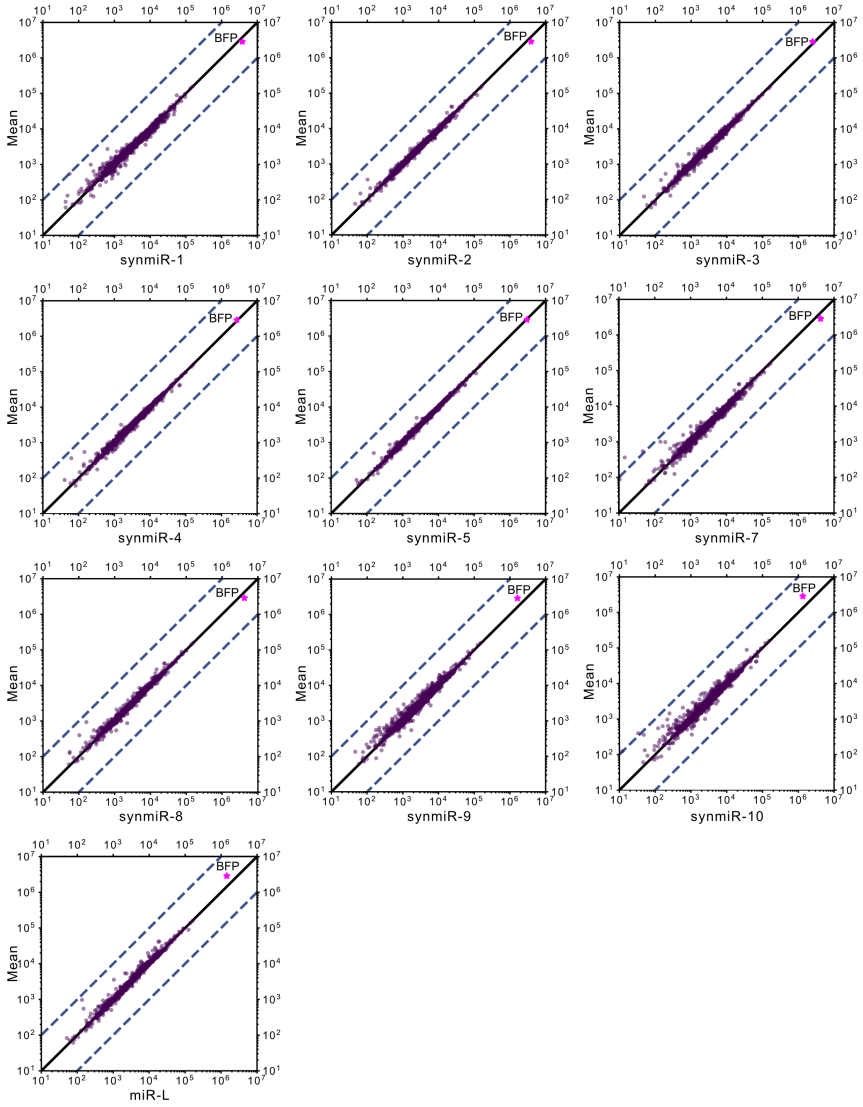

Figure S2.14: Plots of the normalized transcripts per million (TPM) of the synthetic miRNA expressing cells versus the mean TPM. Related to Figure 2.5. The mean TPM was calculated by averaging all the TPM of all the synmiR-expressing cell samples. Solid line indicates where the TPM of the sample is equal to the mean. Dashed lines indicate 10 fold expression differences between the mean and the synmiR-expressing cells.

miRNA	Differentially expressed genes
synmiR-1	HSPA6, PCSK5, HSPA1A, HSPA1B, KRT17, LOXL4, BAG3, SSC4D, FADS2, GNB2, TUBA1A, TNNC1, CLPTM1L, FOXD1, CCN2
synmiR-2	HSPA6, PCSK5, HSPA1A, HSPA1B, RPL17
synmiR-3	HSPA6, PCSK5, HSPA1A, HSPA1B, SPARC
synmiR-4	HSPA6, PCSK5, HSPA1A, HSPA1B, CLPTM1L, POLR2L
synmiR-5	HSPA6, PCSK5, HSPA1A, HSPA1B, CLPTM1L, SSC4D, RPL17, TNNC1
synmiR-7	HSPA6, PCSK5, HSPA1A, HSPA1B, BAG3, GNB2, AP2S1, CLPTM1L, RHOC, HDLBP
synmiR-8	HSPA6, PCSK5, HSPA1A, HSPA1B, RPL17, LOXL4, FADS2, KRT17, RPS2, SSC4D, SPARC, DDX5, TNNC1
synmiR-9	HSPA6, LOXL4, FADS2, KRT17, PCSK5, RPL17, HSPA1B, SSC4D, DKK3, LOXL1, TPM1, TUBA1A, BCAM, SPARC, MAGED1, SEMA6B, TNNC1, AP2S1, RPS2, RPS19
synmiR-10	HSPA6, LOXL4, FADS2, KRT17, PCSK5, GNAS, RPL17, SSC4D, DKK3, BCAM, MAGED1, DDX5, SPARC, TUBA1A, YWHAZ, TNNC1, AP2S1, ACTG1, MRFAP1, RPS2, VIM
miR-L	HSPA6, GNAS, HSPA1B, RPL17, VIM

Table 2.3: Gene annotation for the significantly differentially expressed genes suggested by bulk RNAseq in Data S2-5. Shared differentially expressed genes were color-coded.

Data S2-6

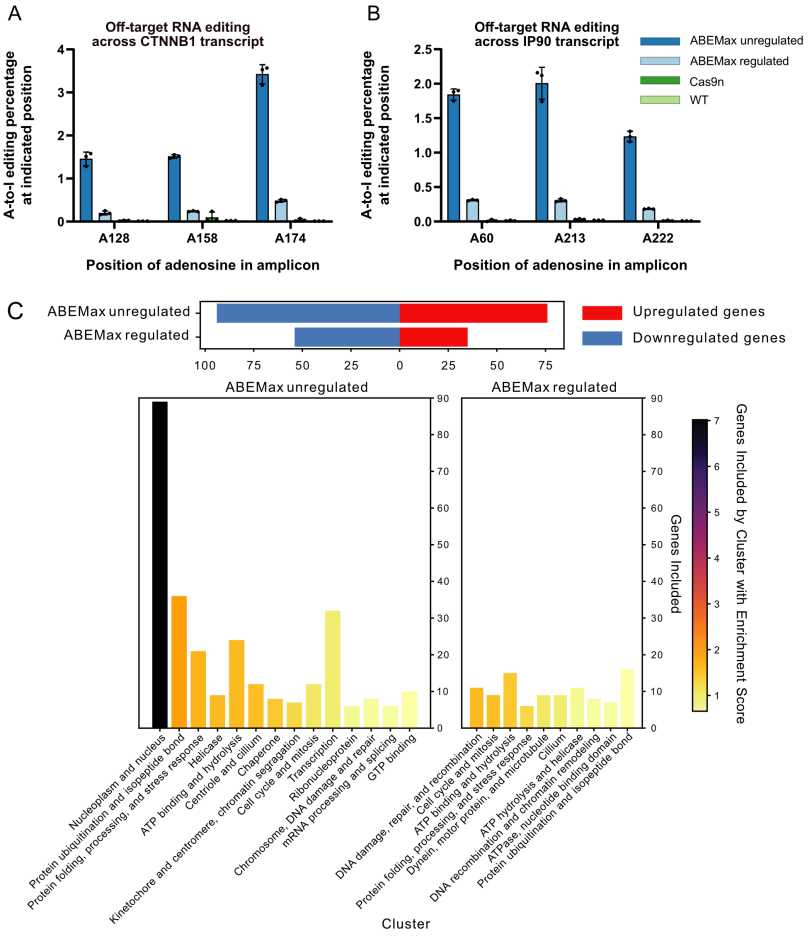

Figure S2.15: DIMMER circuits reduce the off-target RNA editing of the ABEMax base editor. Related to Figure 2.7. The off-target A-to-I RNA editing percentage of the four specific A sites in the CTNNB1 **(A)** and the IP90 **(B)** transcript. Each dot is a biological replicate. Error bars show the standard deviation.
(C) Gene ontology annotation of significantly affected genes identified by bulk RNA sequencing (as in **Figure 2.7F** and **STAR Methods**). Upper panel, the number of significantly affected genes in the unregulated ABEMax and 4×18nt circuit-regulated ABEMax groups. Lower panel, gene ontology annotation of significantly upregulated/downregulated genes in the unregulated ABEMax group (left) and the 4×18nt circuit-regulated ABEMax group (right). Color bar denotes enrichment score.

2.6 STAR Methods

2.6.1 RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Michael B. Elowitz (melowitz@caltech.edu).

Materials availability

Plasmids generated in this study are being submitted to Addgene. All unique/stable reagents generated in this study are available from the lead contact with a completed Materials Transfer Agreement.

Data and code availability

- The original and processed data of this study has been deposited at CaltechDATA (data.caltech.edu) and is publicly available at the time of publication. (DOI: 10.22002/thgnh-6aj57).
- The original code to process data and mathematically model the circuits has been deposited at CaltechDATA (data.caltech.edu, DOI: 10.22002/thgnh-6aj57)
- Any additional information required to reanalyze the data reported in this paper is available from thelead contact upon request.

2.6.2 EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Cell culture

U2OS cells (ATCC Cat# HTB-96), T-Rex cells (ThermoFisher Cat# R71007), CHO cells (ATCC Cat# CCL-61), HEK293 cells (ATCC Cat# CRL-1573), and N2A cells (ATCC Cat# CCL-131) were cultured at 37 °C in a humidity-controlled chamber with 5% CO2. The growth media consisted of DMEM (Dulbecco's Modified Eagle Medium, Thermo Fisher #11960-069) supplemented with 10% FBS, 1 U/ml penicillin, 1 µg/ml streptomycin, 1 mM sodium pyruvate, 1×NEAA (Thermo Fisher #11140-050), 1 mM L-glutamine, and 0.1 mg/mL Normocin (InvivoGen #ant-nr).

Animals

All animals used in this study were approved by the Institutional Animal Care and Use Committee (IACUC) at the California Institute of Technology. Wildtype C57BL female mice (1.5-4 months of age) were ordered from Jackson Laboratories (Bar Harbor, ME, USA) and were used for surgeries, viral injection, and head-fixed imaging. All mice were singly-housed post-operation and for the remainder of the data acquisition in a room with a reverse light cycle (12h – 12h). For all imaging sessions, mice were imaged during the dark phase of their light cycle.

2.6.3 METHOD DETAILS

Plasmid construction

Constructs used in this study are listed in Key Resource Table. Some constructs were generated using standard cloning procedures. The inserts were generated using PCR or gBlock synthesis (IDT) and were ligated either by T4 ligase (NEB #M0202M) or In-Fusion (Takara #102518) assembly with backbones that are linearized using restriction digestion. *E.coli* competent cells (NEB Cat# C3040H and NEB Cat# C3019H) were utilized to amplify the plasmids. The rest of the constructs were designed by the authors and synthesized by GenScript. Selected constructs will be deposited at Addgene and the maps are available.

miRNA alignment to the database

Each synthetic miRNA sequence (mature miRNA, 22 nt) is aligned to the known miRNA sequence database [76, 77, 78, 79, 80, 81, 82, 83] (https://mirbase.org/) to identify if there are any similarities existing between the synthetic sequences and the natural sequences.

Transient transfection

Cells were seeded at a density of 50,000 cells in each well of a 24 well plate (or at a density of 10,000 cells in each well of a 96 well plate), either standard for flow cytometry or glass-bottom for imaging experiments, and cultured under standard conditions overnight. The following day, the cells were transiently transfected using Fugene HD (Promega #E2311), according to the manufacturer's protocol.

Flow cytometry

Cells were incubated 2 days after transient transfection, and the culture media was replaced 24 hours post-transfection. Cells were trypsinized with 75 µL of 0.25% trypsin for 5 minutes at 37 °C. After digestion, cells were resuspended with 125 µL of HBSS containing 2.5 mg/ml BSA and 1 mM EDTA. Cells were then filtered through a 40 µm cell strainer and analyzed using a CytoFLEX S instrument (Beckman Coulter).

Cell sorting

To prepare the mono-clones that expressed the genomic-integrated DIMMER circuit, cells were harvested and resuspended in sorting buffer (BD FACS Pre-Sort Buffer) supplemented with 1 U/ml DNase I by the cell sorter (Sony MA900) as mono-clones. Cells were sorted into 96 well plates in the normal U2OS culture media. Cells were expanded in the 24 well plate before flow cytometry measurement.

Bulk RNA sequencing to identify the off-target effects of the synthetic miRNA on the transcriptome

- Sample preparation and sequencing

To verify the off-target effect of all the synmiRs, U2OS cells were plated on 6-well plates with 300,000 cells per well. Cells were transfected the following day with 1,000 ng of either the control plasmid or the BFP-miRNA plasmid using Fugene HD (Promega #E2311) according to the manufacturer's instructions. Media was replaced with 2 mL of fresh media 24 hours post-transfection. Cells were harvested 48 hours post-transfection by digestion with 0.25% Trypsin-EDTA, centrifugation at 300g for 5 minutes, and removal of the supernatant by aspiration. The cell's pellet was stored in -80 °C prior to the purification.

RNA was extracted using the RNeasy kit (Qiagen #74106) according to the manufacturer's instructions. RNA was treated with Turbo DNase (Thermo Fisher #AM2238) and purified using the RNeasy kit RNA cleanup protocol. mRNA sequencing libraries were prepared by Novogene.

DNA-PAINT

- Buffers

The following buffers were used:

- NH_4Cl solution: NH_4Cl (Roth, no. K298.1) was dissolved in ddH_2O for a 2 M stock solution, filtered with 0.2 μm filter.

- Blocking buffer: 1×PBS, 1 mM EDTA (Thermo Fisher, no. AM9260G), 0.02% Tween-20 (Life Science, no. P7949), 0.05% NaN_3 (Serva, no. 30175.01), 2% BSA (Sigma-Aldrich, no. A9647-100G), 0.05 mg/ml sheared salmon sperm DNA (Life Technologies, no. 15632011), filtered with 0.2 μm filter.

- Imaging buffer: 1×PBS, 1 mM EDTA, 0.02% Tween-20, 500 mM NaCl (Thermo Fisher, no. AM9760G), supplemented with PCA, PCD, trolox, filtered with 0.2 µm filter.

- PCA, PCD and Trolox

- Trolox (100×) was prepared by dissolving 100 mg (±)-6-hydroxy-2,5,7,8-tetra-methylchromane-2-carboxylic acid (trolox; Sigma-Aldrich, 238813-5G) in 430 µL of 100% methanol (Sigma-Aldrich, 32213-2.5L), 345 µL of 1 M NaOH (VWR, 31627.290) and 3.2 mL of water.

- PCD (40×) was made by mixing 154 mg of 3,4-dihydroxybenzoic acid (PCA; Sigma-Aldrich, 37580-25G-F) in 10 mL of water and NaOH and adjusting the pH to 9.0. PCD (100×) was prepared by adding 9.3 mg of protocatechuate 3,4-dioxygenase pseudomonas (PCD; Sigma-Aldrich, P8279) to 13.3 mL of buffer containing 100 mM Tris-HCl pH 8.0 (Thermo Fisher Scientific, AM9855G), 50 mM KCl (Thermo Fisher Scientific, AM9640G), 1 mM EDTA, and 50% glycerol (Sigma-Aldrich, 65516-500ml)).

- Cloning

An mEGFP gBlock (obtained from IDT) was inserted into a pcDNA3.1(+) backbone (Thermo Fisher, no.V79020) via Gibson assembly. Two codon-optimized fragments of human EGFR (obtained from IDT) were fused to the mEGFP-pcDNA3.1(+) backbone via Gibson assembly. The plasmid concentration was measured with the NanoDrop One (Thermo Fisher Scientific).

- Cell culture

CHO-K1 cells (ATCC: CCL-61) were cultured in Ham's F-12K (Kaighn's) medium (Gibco, no. 21127022) supplemented with 10% FBS (Gibco, no. 11573397). U2OS-CRISPR-Nup96-mEGFP cells (a gift from the Ries and Ellenberg laboratories) were cultured in McCoy's 5A medium (Gibco, no. 16600082) supplemented with 10% FBS. All cells were cultured at 37 °C and 5% CO2 and split every 2-3 days via trypsinization using trypsin-EDTA (Gibco, no. 25300096).

- Nanobody-DNA conjugation

First, the anti-GFP nanobody (clone 1H1, Nanotag Biotechnologies, N0305) and anti-rabbit IgG nanobody (Nanotag Biotechnologies, N2405) were conjugated to

a DBCO-PEG4-Maleimide linker (Jena Bioscience, no. CLK-A108P). After removing the unreacted linker with Amicon centrifugal filters (10,000 MWCO), the DBCO-nanobody was conjugated via DBCO-azide click chemistry to the docking strand (Metabion, see sequence in Table S2.4). A detailed description of the conjugation can be found in the former work [84].

- Fixation of cells

The cells were fixed with 37 °C pre-warmed methanol-free 4% PFA (Thermo Fisher, no. 043368.9M) in 1×PBS for 15 min. Then, the cells were washed 3 times with 1×PBS and then permeabilized with 0.125% TritonX-100 (Sigma-Aldrich, no. 93443) in 1×PBS for 2 min. After washing 3 times with 1×PBS, the cells were blocked with the blocking buffer either overnight or for at least 3h at 4 °C.

- Sample preparation of U2OS cells

10,000 cm^{-2} U2OS cells were seeded on an ibidi eight-well high glass-bottom chambers (no. 80807). On the next day, the cells were fixed as stated in the fixation protocol and blocked for 3h at 4 °C. After washing 3 times with 1×PBS, the primary EGFR antibody (Cell Signaling, clone D38B1, no. 4267) with a dilution of 1:200 in blocking buffer was incubated overnight at 4 °C. The next day, the sample was washed 3 times with 1×PBS and 25 nM of R2 anti-rabbit NB in the blocking buffer was incubated at RT for 1h. After washing 3 times with 1×PBS, the sample was post-fixed with 4% PFA in 1×PBS for 5 min at RT. The cells were then quenched with 200 mM NH_4Cl (Roth, no. K298.1) for 5 min and washed 3 times with 1×PBS. 90nm gold-nanoparticles (Absource, no. G-90-100) in 1:1 in 1×PBS were incubated for 5 min at RT. After washing 3 times with 1×PBS, the cells were washed once with the imaging buffer.

- Sample preparation for plasmid dosage measurement

5,000 cm^{-2} CHO-K1 cells were seeded on an ibidi eight-well high glass-bottom chambers (no. 80807) one day before transfection. The cells were transfected with EGFR-mEGFP plasmids with a Thermo Fisher Lipofectamine 3,000 reagent (no. L3000008) with the lower Lipofectamine concentration as indicated by the manufacturer and different plasmid concentrations (0 ng, 125 ng, 250 ng and 500 ng of plasmid per well). After 48h of transfection, the cells were fixed as indicated in the fixation protocol and blocked with blocking buffer for 3h at 4 °C. The cells were then post-fixed with 4% PFA and 0.2% glutaraldehyde (Serva, no. 23115.01) in 1×PBS for 10 min. After quenching the sample with 200 mM NH_4Cl for 5 min and washing 3 times with 1×PBS. 90 nm gold-nanoparticles in 1:1 in 1×PBS were

incubated for 5 min at RT. After washing 3 times with 1×PBS, the cells were washed once with the imaging buffer.

- Sample preparation of DIMMER plasmids

5,000 cm^{-2} CHO-K1 cells were seeded on an ibidi eight-well high glass-bottom chambers (no. 80807) one day before transfection. The cells were transfected with EGFR-mEGFP plasmids with a Thermo Fisher Lipofectamine 3,000 reagent (no. L3000008) with the lower Lipofectamine concentration as indicated by the manufacturer and 250 ng plasmid per well (200 µL solution per well and 25 µL transfection solution). After 48h of transfection, the cells were fixed as indicated in the fixation protocol and blocked with the blocking buffer for 3h at room temperature (RT). 25 nM R3 anti-GFP nanobodies were incubated in the blocking buffer for 1h at RT. After washing 3 times with 1×PBS, the nanobodies were post-fixed with 4% PFA and 0.2% glutaraldehyde in 1×PBS for 10 min. The cells were then quenched with 200 mM NH_4Cl for 5 min and washed 3 times with 1×PBS. 90nm gold-nanoparticles (Absource, no. G-90-100) in 1:1 in 1×PBS were incubated for 5 min at RT. After washing 3 times with 1×PBS, the cells were washed once with the imaging buffer and imaged in the imaging buffer.

- DNA-PAINT imaging

The samples were imaged in the imaging buffer with the corresponding imager strand (obtained from Metabion, see table below for imager strand sequences) for 40k frames with 100ms exposure time per frame and a readout rate of 200 MHz.

Sequence name	**Docking strand sequence (5′ to 3′)**	**Imager strand sequence (5′ to 3′)**
R2	ACCACCACCACCACCACCA	TGGTGGT-Cy3B
R3	CTCTCTCTCTCTCTCTCTC	GAGAGAG-Cy3B

Table 2.4: DNA sequence of docking strands and imager strands.

- Microscope setup

The samples were measured on inverted total internal reflection fluorescence (TIRF) microscopes (Nikon Instruments, Eclipse Ti2) which are equipped with an oil-immersion objective (Nikon Instruments, Apo SR TIRF ×100/numerical aperture 1.49, oil) and a perfect focusing system. The mRuby3 signal was bleached by the 560 nm laser (MPB Communications, 1 W) by using Highly inclined and laminated optical sheet (HILO) illumination. Afterwards, the TIRF mode was established.

The Cy3B-conjugated imagers were excited with the 560 nm laser. The laser beam was cleaned with a filter (Chroma Technology, no. ZET561/10) and coupled into the microscope with a beam splitter (Chroma Technology, no. ZT561rdc). The fluorescent signal was filtered with an emission filter (Chroma Technology, nos. ET600/50m and ET575lp) and projected onto a sCMOS camera (Hamamatsu Fusion BT) without further magnification. The camera's central 1152×1152 pixels (576×576 pixels after binning) were used as the region of interest, with a resulting effective pixel size of 130nm. The raw microscopy data was acquired via μManager (Version 2.0.1).

Base editor experiment

- Transfection and Sample collection

HEK293 cells were seeded at the density of 150k/well in the 24-well plate one day before the transfection. For each well, 750 ng ABEMax base editor plasmid and 250 ng sgRNA plasmid were co-transfected using Fugene HD (Promega #E2311), according to the manufacturer's protocol. Cells were harvested 72 h post-transfection.

- Amplicon sequencing

To quantify the on-target base editor editing rate, cells were sorted based on the constitutively-expressing mRuby on the base editor plasmid and EGFP on the sgRNA plasmid. Genomic DNA was extracted using Qiagen DNeasy kit (cat. nos. 69504) according to the manufacturer's protocol. The genomic DNA amplicons were amplified using the primer in the reference [47], with the adaptor sequence on each end. Amplicons were size-verified by DNA electrophoresis, purified by the Qiagen gel purification kit, and sequencing was performed by Genewiz Amplicon-EZ (150-500bp) service.

To quantify the off-target base editor editing rate on the genome, genomic DNA was extracted, and amplicons were obtained using the potential genomic DNA off-target loci-targeted primers reported previously [47]. To quantify the off-target editing on the RNA level, the total RNA was extracted from the cell using the Zymo Direct-zol RNA miniprep kit, according to the manufacturer's protocol. cDNA samples were prepared using the Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). Amplicons were obtained using the potential RNA off-target-loci-targeted primers reported previously [47]. The amplicons were then size-verified by DNA electrophoresis, purified with the Qiagen gel purification kit,

barcoded with the Illumina Miseq 16S Metagenomic sequencing index primers, and further purified with the magnetic NGS beads (Omega Bio-Tek). Library was quantified and normalized with the Qubit fluorometer. Library was then denatured and sequenced with Element AVITI System Sequencing Instrument using AVITI 2×150 Sequencing Kit Cloudbreak (Catalog # 860-00013).

- Bulk RNA sequencing

To analyze the transcriptome-wide perturbation of the base editor, total RNA was extracted from the cells with the Zymo Direct-zol RNA miniprep kit (lot # R2050). 50 ng of extracted mRNA from each sample were used as inputs for downstream NGS library preparation.

mRNA-seq libraries were prepared in 96-well format with a modified 3'Pool-seq protocol [85]. In brief, reverse transcription reaction were prepared by mixing input RNA with 1 µl Indexed RT Primer (10 µM), 1 µl 10 mM dNTP Mix (New England Biolabs Cat# N0447S),1 µl diluted ERCC Spike-In Mix 1 (0.004 µL stock ERCC per µg RNA, Thermo Fisher Cat# 4456740), 3.6 µl of 5×RT buffer (Thermo Fisher Cat# EP0752), 0.5 µl of RNAse inhibitor (Thermo Fisher Cat# EO0381), 1 µl Maxima RT H minus (Thermo Fisher Cat# EP0752), 2.5 µl 10 uM Template Switching Oligo into a 18 µl reaction. Reverse transcription was carried out in a thermocycler with a program described in 3'Pool-seq protocol.

Samples from each row of 96-well plate were pooled (column pooling) by mixing an equal volume of each Reverse Transcription reaction into a new well at a total volume of 20 µl. Residual primers were then degraded with the addition of 1 µl Exonuclease I (New England Biolabs) and incubated at 37 °C for 45 min followed by denaturation at 92 °C for 15 min. Subsequent cDNA amplification, tagmentation, and row pooling was performed following 3'Pool-seq protocol.

Finally, 20 µl of pooled NGS library were subject to Gel-based size selection using E-Gel EX Agarose Gel (Thermo Fisher Cat# G401001) to enrich for fragments with size range between 200-1000 bp and eluted in 15 µl.

Eluted pooled NGS libraries were examined in an Agilent TapeStation 4200 (Agilent Technologies) to determine average fragment sizes. Library concentration was quantified in a Qubit 3.0 Fluorometer (Life Technologies). NGS library molarity was then calculated using 660 g/mol per base-pair as a molecular weight. NGS libraries were diluted to 2 nM, denatured in 0.2 N NaOH, and loaded onto Element AVITI sequencer following Element Biosciences Cloudbreak Sequencing user guide.

DNA-FISH

To cross-validate the success of DIMMER-regulating dCas9 labeling the telomeres, Nikon 2 Multipoint acquisition was used to record the absolute coordinates of the fields of view. After the live cell imaging, the plate was removed from the scope and DNA-FISH was performed. Dye-conjugated probes designed to bind to the distal telomeric regions of the chromosome were used following the previously reported protocol [86, 87]. The plate was then loaded on the scope, and the recorded coordinates were retrieved. The exactly same fields of view were matched by eye around the recorded coordinates.

DIMMER circuit dynamical quantitation in mice cranial window

- AAV Viral Production

All constructs were cloned using the cloning strategy described above. The viral titers of the viruses used in this study were given in the table below. CAP-B22 and CAP-B10 are AAV capsids developed in the previous study [51]. The unit of the viral titers is viral genome copies per mL.

Group	**Cargo**	**Viral Titers**
CaMKIIα_unregulated	CAP-B10.CaMKIIα.H2B.EGFP.sv40pA	1.06×10^{14}
	CAP-B10.CaMKIIα.H2B.mRuby3.sv40pA	2.12×10^{14}
CaMKIIα_regulated	CAP-B10.CaMKIIα.H2B.EGFP(4×19nt regulated).sv40pA	8.11×10^{13}
	CAP-B10.CaMKIIα.H2B.mRuby3.sv40pA	2.12×10^{14}
MeP_unregulated	CAP-B22.CMVenhancer.MEP229.H2B.EGFP.sv40pA	1.22×10^{14}
		4.04×10^{13}
	CAP-B22.CMVenhancer.MEP229.H2B.mRuby3.sv40pA	1.94×10^{14}
MeP_regulated	CAP-B22.CMVenhancer.MEP229.H2B.EGFP(4×19nt regulated).sv40pA	4.14×10^{14}
		9.86×10^{13}
	CAP-B22.CMVenhancer.MEP229.H2B.mRuby3.sv40pA	1.94×10^{14}
mDLX2_unregulated	CAP-B10.mDLX2enhancer.H2B.EGFP.sv40pA	1.53×10^{14}
	CAP-B10.mDLX2enhancer.H2B.mRuby3.sv40pA	1.64×10^{14}
mDLX2_regulated	CAP-B10.mDLX2enhancer.H2B.EGFP(4×19nt regulated).sv40pA	7.79×10^{13}
		4.98×10^{13}
	CAP-B10.mDLX2enhancer.H2B.mRuby3.sv40pA	1.64×10^{14}

Table 2.5: The viral titers used in the mice cranial window study. The unit of the viral titers is viral genome copies per mL.

- Cranial Window Surgeries

Mice were deeply anesthetized under 1-2% isoflurane and given subcutaneous injections of ketofen (5 mg/kg), buprenorphine (3.25 mg/kg), dexamethasone (2 mg/kg) and saline for analgesia and hydration. A subcutaneous injection of a few drops of bupivacaine was given under the incision site for local analgesia. An incision was performed over the skull and a piece of skin near the left motor cortex was removed to make room for the implant. A circular craniotomy of approximately 3-mm in diameter was drilled over the left motor cortex and centered 1.6-mm anterior and 1.6-mm lateral from bregma. Saline was flushed periodically throughout the drilling over the bone to hydrate and prevent heating. After extraction of the cranial bone, a gelatin sponge soaked in saline was applied over the exposed dura to remove debris and stop any residual bleeding. A small drop of Kwik-sil silicone elastomer was applied over the dura and a 3-mm diameter glass coverslip (#1.5) was implanted over the craniotomy and sealed with dental cement. A custom-made titanium headbar was fixed using cement over the window to allow for later head-fixation during 2-photon imaging. Following surgery, mice were allowed to recover for at least 4-days before beginning habituation to a rotary platter setup.

- AAV Vectors and Administration in Mice

AAV packaging, purification, and delivery were performed as described in the earlier study [88]. Following surgical implants of cranial windows and recovery, mice were anesthetized under 1-2% isoflurane. Intravenous administration of AAVs was performed by injection into the retro-orbital sinus of mice. All vectors were each delivered at a total of 5×10^{12} viral genomes (VG) in a volume of 50 µL per mouse from 2-4 months of age. A total of three cohorts were designed to test for the performance of the DIMMER circuits with different gene-regulatory elements: CMV-enhancer paired with a minimal MeCP2 promoter (CMV enhancer-MEP229), an mDLX2 enhancer paired with a minimal beta-globin promoter (mDLX2-minBglobin), and a CaMKIIα promoter. Each cohort had two groups: One group co-expressing an unregulated H2B-EGFP transgene and an unregulated H2B-mRuby3 transgene, and a regulated group co-expressing a 4×19nt DIMMER regulated H2B-EGFP transgene and an unregulated H2B-mRuby3 transgene. These transgenes were co-delivered separately in their own AAV-capsids. The mRuby3 signal was later used as an internal reference for the EGFP signal for each mouse.

- Two-Photon Head-Fixed Imaging and Acquisition

Mice were habituated to the handler's hand for 3 days. Following this, they were head-fixed onto a rotary platter for an additional 3 days to further habituate to the

imaging setup. After habituation, mice were imaged over a variable number of days and weeks to capture the progression of EGFP and mRuby3 expression. We used a custom home built 2-photon microscope equipped with a galvo-galvo scanner for imaging. Each imaging session involved acquiring two high-resolution Z-stacks per mouse using a 20×/1.0 NA objective (Zeiss W Plan-Apochromat, Cat. No. 421452-9700-000) with water immersion. Each of these Z-stacks consisted of 10 Z-steps, ranging from 200-250 µm to 300-350 µm below the dura, with each Z-step separated by 10 µm intervals, for a final 100 µm thick Z-stack. Each Z-step was acquired at a scan rate of 1 Hz with a resolution of 256 by 256 pixels and a pixel dwell time of 13 µs for a total of 180 seconds, over an area of 400 µm × 400 µm. Z-steps in each Z-stack were later processed as mean-intensity projected TIF files to generate a single Z-step image. Each Z-stack was first acquired at 940 nm wavelength (EGFP), followed by 1100 nm (mRuby3). The average laser power beneath the objective was maintained across imaging sessions at 36-37 mW for the 940 nm wavelength and 46-47 mW for the 1100 nm wavelength.

2.6.4 QUANTIFICATION AND STATISTICAL ANALYSIS

Flow cytometry data analysis

We used the FlowJo V10 and self-built python code to analyze the flow data. The regression curves and the confidence intervals in **Figure 2.5B** and **Figure S2.4** were computed using the statsmodels.regression.linear_model.OLS package in logarithm spaces. The statistical details of each experiment are described in the results and figure legends.

In-vitro image analysis

The transiently transfected U2OS/HEK293 cells were imaged using a Nikon confocal microscope at 60× magnification, such that each image was spaced by 0.5 microns in the z-direction. For HEK293 cells, 1 slice in the z-direction was taken given the relatively smaller volume compared to U2OS. Images were processed by the Fiji software [89].

To analyze the relative signal intensity in the dCas9 imaging experiment, maximum intensity projection of 11 slices of the z-stacks were applied. To determine the signal intensity of the dots, freehand lines were drawn to select the dot regions. To determine the signal intensity in the background, $\tilde{5}$ micron-long straight lines centering the dots were drawn. The signal intensities were generated by the Fiji ROI

mean intensity function. The relative signal intensity is calculated by normalizing the background intensity to be 1.

To analyze the signal to noise ratio (SNR), freehand lines were drawn to select the dot regions and the nucleus regions. The noise intensity was calculated by the intensity in the nucleus excluding the dots area. The SNR was calculated dividing the mean intensities of each dot by the noise.

Bulk RNA sequencing data analysis to identify the off-target effect of the synthetic miRNAs

- Preprocessing of sequencing data

Reads from the RNA sequencing were aligned to a custom reference genome using kallisto (0.48.0) [90]. This reference consisted of the human genome GRCh38 cDNA (https://github.com/pachterlab/kallisto-transcriptome-indices/releases) and mTagBFP2 coding sequences. Weakly expressed genes were filtered out if they exhibited fewer than 3 samples expressing at least 10 transcripts per million (TPM), or if the maximum TPM among all samples was less than 105. Then, filtered counts were input to DEseq to eliminate the impact of size factors. As the BFP-only cells were used as a reference to evaluate the off-target effect of the miRNAs, genes that showed fluctuating expressions among the three biological replicates of BFP-only cells should be removed from analysis. To achieve this, we computed $\log(1 + x)$ where x denotes the normalized TPM among the three biological replicates of BFP-only cells. The Fano factors [91, 92] of these logarithmic expressions were determined and ranked. Transcripts that ranked as the largest 2.5% in logarithmic Fano factors were eliminated from further analysis. Finally, we computed $\log(1 + x)$, where x denotes the normalized TPM among all the samples. A difference function was defined to compute the absolute value of the $\log(1 + x)$ difference between each sample and the untransfected sample. The medians of the difference function of the BFP-only groups and the experimental groups were calculated and used for comparison. The difference between those two difference functions were ranked and similarly, transcripts that ranked as the largest 3% were removed from further analysis.

The Fano factor is defined as:

$$\text{Fano factor} = \frac{\text{variance}(\log(1+x))}{\text{mean}(\log(1+x))}$$

The equation of the difference function is defined as:

$$\Delta(\text{BFP}, \text{untransfected}) = |\log(1 + x_{\text{BFP}}) - \log(1 + x_{\text{untransfected}})|$$

$$\Delta(\text{experimental}, \text{untransfected}) = |\log(1 + x_{\text{experimental}}) - \log(1 + x_{\text{untransfected}})|$$

$\Delta_{\text{ranked}} = \text{median}(\Delta(\text{BFP}, \text{untransfected})) - \text{median}(\Delta(\text{experimental}, \text{untransfected}))$

- Differential gene expression analysis

To characterize the perturbations that synthetic miRNA brought to the endogenous transcriptome, differential expression analysis was performed using DESeq2 (1.40.1) [93] in R (4.3.1) comparing transcript counts in miRNA transfected cells and BFP-only cells.

DNA-PAINT data analysis

Obtained fluorescent data was reconstructed with Picasso software [35]. The data was first drift-corrected with redundant cross-correlation, after that with picked gold particles as fiducials.

In order to determine the receptor density, a homogeneous area of the cells was picked and the DNA-PAINT data was clustered with the SMLMS clustering algorithm of Picasso [35, 94].

The identified cluster centers were used to calculate the measured receptor density per μm^2 (number of cluster centers per area). Given that the labeling efficiency (LE) of the binders to their targets was less than 100%, a correction factor was applied to account for incomplete labeling. The EGFR density was calculated by multiplying the measured receptor density with the LE of the respective binders. The LE values of the binders were determined as previously described in the former work [95]. Specifically, the LE of the anti-GFP nanobody was determined to be 37%, while the LE of the EGFR antibody was 71%.

To test the statistical significance of the determined receptor densities, we performed the Mood's median test. Mood's median test is a non-parametric statistical test which tests whether the median of two groups are statistically different. Our statistical tests were performed with a custom python script using functions from the scipy.stats module [96].

Base editor experiment data analysis

- Amplicon sequencing data analysis

The on-target and off-target editing rates were analyzed using the online tool CRISPResso2 [97].

- Bulk RNA sequencing data analysis

Weakly expressed genes were filtered out if they exhibited fewer than 3 samples expressing at least 10 transcripts per million (TPM). Differential expression analysis was performed using DESeq2 (1.40.1) [93] in R (4.3.1). Identified significantly differentially expressed genes were eliminated from further analysis if they were identified among the unregulated, regulated, and untransfected groups with the same trend, which indicates that these genes tend to be expressed as more similar profiles among the base editor-transfected groups and the WT group, thus very likely to be the artifact of the filler-plasmid expression. GO annotation was performed using the DAVID web server (https://davidbioinformatics.nih.gov/).

Two photon microscopy data analysis

Two-photon Z-stacks were acquired separately for EGFP and mRuby3 channels and processed independently. Raw data files were segmented into individual z-step movies based on predefined frame intervals using a custom MATLAB script, with each segment saved as a multi-page TIFF file [98]. Motion correction was applied to each TIFF using rigid registration, and motion-corrected data were used to generate mean- and maximum-intensity projection images for each z-step. Following motion correction, mean-intensity projection images were flatfield corrected using MATLAB's imflatfield function with a specified radius of 20 pixels to remove uneven illumination artifacts. Corrected images were saved in a dedicated output directory for subsequent analysis. To ensure spatial correspondence between channels, each flatfield-corrected mRuby3 image was aligned to its corresponding EGFP image using normalized cross-correlation, correcting for lateral drift and maximizing overlap between channels.

Cell segmentation was performed on the mRuby3 channel using the Cellpose [99] deep learning-based segmentation algorithm (model type = "cyto", diameter = 9.0 pixels), generating individual cell masks for each z-step image. These masks were saved in both TIFF and NumPy formats and served as a consistent reference for fluorescence quantification across channels. For each z-step, the corresponding mRuby3-derived cell mask was applied to the EGFP image to extract the total EGFP fluorescence per cell by summing the pixel intensities within each ROI. Separately, a dedicated script was used to quantify total mRuby3 fluorescence per cell by applying the same cell mask to the original mRuby3 image for that z-step. Results were saved as per-z-step .csv files, with all individual pixel values archived in cumulative text files for downstream statistical analysis and visualization.

To evaluate single-cell expression and dynamics across time, total EGFP and mRuby3 fluorescence values per cell were aggregated across multiple mice and imaging timepoints. For the CaMKIIα and CMVe-MeCP2 cohorts, z-steps 1, 5, and 10 were selected from each 10-step Z-stack to reduce the likelihood of overcounting the same cells across adjacent planes. For the mDLX2-minβglobin cohort, only z-step 5 was selected.

2.7 Online Supplemental Materials

Supplemental materials below are available online as separate files at:

Key Resource Table

Table S1, List of miRNA and targets used in this study, related to Figure 2.1-2.7

Table S2, Gene Ontology annotations for significantly regulated genes in ABEMax unregulated vs. mock and ABEMax regulated vs. mock, related to Figure 2.7

References

[1] Michael J Flynn et al. "Synthetic dosage-compensating miRNA circuits allow precision gene therapy for Rett syndrome". In: *bioRxiv* (Mar. 2024).

[2] Kate Franz, Abhyudai Singh, and Leor S Weinberger. "Lentiviral vectors to study stochastic noise in gene expression". In: *Methods Enzymol.* 497 (2011), pp. 603–622.

[3] Sowmya Balasubramanian et al. "Rapid recombinant protein production from piggyBac transposon-mediated stable CHO cell pools". In: *J. Biotechnol.* 200 (Apr. 2015), pp. 61–69.

[4] Zakary S Singer et al. "Dynamic heterogeneity and DNA methylation in embryonic stem cells". In: *Mol. Cell* 55.2 (July 2014), pp. 319–331.

[5] Hédia Maamar, Arjun Raj, and David Dubnau. "Noise in gene expression determines cell fate in Bacillus subtilis". In: *Science* 317.5837 (July 2007), pp. 526–529.

[6] Uri Alon. "Network motifs: theory and experimental approaches". In: *Nat. Rev. Genet.* 8.6 (June 2007), pp. 450–461.

[7] Leonidas Bleris et al. *Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template*. Issue: 1 Pages: 519 Publication Title: Molecular Systems Biology Volume: 7. 2011.

[8] Timothy J Strovas et al. "MicroRNA-based single-gene circuits buffer protein synthesis rates against perturbations". In: *ACS Synth. Biol.* 3.5 (May 2014), pp. 324–331.

[9] Jin Yang et al. “A synthetic circuit for buffering gene dosage variation between individual mammalian cells”. In: *Nat. Commun.* 12.1 (July 2021), p. 4132.

[10] Kasey S Love et al. “Model-guided design of microRNA-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells”. In: *bioRxiv* (July 2024).

[11] Jessica Sheu-Gruttadauria and Ian J MacRae. “Phase Transitions in the Assembly and Function of Human miRISC”. In: *Cell* 173.4 (May 2018), 946–957.e16.

[12] David P Bartel. “Metazoan MicroRNAs”. In: *Cell* 173.1 (Mar. 2018), pp. 20–51.

[13] Christof Fellmann et al. “An optimized microRNA backbone for effective single-copy RNAi”. In: *Cell Rep* 5.6 (Dec. 2013), pp. 1704–1713.

[14] Linghua Qiu et al. “A construct with fluorescent indicators for conditional expression of miRNA”. In: *BMC Biotechnol.* 8 (Oct. 2008), p. 77.

[15] Elad Elkayam et al. “Multivalent Recruitment of Human Argonaute by GW182”. In: *Mol Cell* 67.4 (Aug. 2017), 646–658.e3.

[16] Gaspare La Rocca et al. “Inducible and reversible inhibition of miRNA-mediated gene repression in vivo”. In: *Elife* 10 (Aug. 2021).

[17] Chanseok Shin et al. “Expanding the microRNA targeting code: functional sites with centered pairing”. In: *Mol Cell* 38.6 (June 2010), pp. 789–802.

[18] David Jee et al. “Dual Strategies for Argonaute2-Mediated Biogenesis of Erythroid miRNAs Underlie Conserved Requirements for Slicing in Mammals”. In: *Mol Cell* 69.2 (Jan. 2018), 265–278.e6.

[19] Peter Y Wang and David P Bartel. “The guide-RNA sequence dictates the slicing kinetics and conformational dynamics of the Argonaute silencing complex”. In: *Mol Cell* 84.15 (Aug. 2024), 2918–2934.e11.

[20] Sean E McGeary et al. “MicroRNA 3’-compensatory pairing occurs through two binding modes, with affinity shaped by nucleotide identity and position”. In: *Elife* 11 (Feb. 2022).

[21] Winston R. Becker et al. “High-throughput analysis reveals rules for target RNA binding and cleavage by AGO2”. In: *Molecular Cell* 75.4 (2019), pp. 741–755.

[22] Yale S Michaels et al. “Addendum: Precise tuning of gene expression levels in mammalian cells”. In: *Nat. Commun.* 10.1 (June 2019), p. 2622.

[23] Stefan L Ameres et al. “Target RNA-directed trimming and tailing of small silencing RNAs”. In: *Science* 328.5985 (June 2010), pp. 1534–1539.

[24] Minju Ha and V Narry Kim. “Regulation of microRNA biogenesis”. In: *Nat. Rev. Mol. Cell Biol.* 15.8 (July 2014). Publisher: Nature Publishing Group, pp. 509–524.

[25] Sean E McGeary et al. “The biochemical basis of microRNA targeting efficacy”. In: *Science* 366.6472 (Dec. 2019).

[26] J Pontén and E Saksela. “Two established in vitro cell lines from human mesenchymal tumours”. In: *Int. J. Cancer* 2.5 (Sept. 1967), pp. 434–447.

[27] T T Puck, S J Cieciura, and A Robinson. “Genetics of somatic mammalian cells. III. Long-term cultivation of euploid cells from human and animal subjects”. In: *J. Exp. Med.* 108.6 (Dec. 1958), pp. 945–956.

[28] F L Graham et al. “Characteristics of a human cell line transformed by DNA from human adenovirus type 5”. In: *J. Gen. Virol.* 36.1 (July 1977), pp. 59–74.

[29] J B Olmsted et al. “Isolation of microtubule protein from cultured mouse neuroblastoma cells”. In: *Proc. Natl. Acad. Sci. U. S. A.* 65.1 (Jan. 1970), pp. 129–136.

[30] Julia M Marchingo and Doreen A Cantrell. “Protein synthesis, degradation, and energy metabolism in T cell immunity”. In: *Cell. Mol. Immunol.* 19.3 (Mar. 2022), pp. 303–315.

[31] Eszter Lakatos et al. “Protein degradation rate is the dominant mechanism accounting for the differences in protein abundance of basal p53 in a human breast and colorectal cancer cell line”. In: *PLoS One* 12.5 (May 2017), e0177336.

[32] Shixuan Liu et al. “Large cells activate global protein degradation to maintain cell size homeostasis”. In: *bioRxiv* (Nov. 2021), p. 2021.11.09.467936.

[33] Kyle Swovick et al. “Interspecies Differences in Proteome Turnover Kinetics Are Correlated With Life Spans and Energetic Demands”. In: *Mol. Cell. Proteomics* 20 (Jan. 2021), p. 100041.

[34] “Direct Readout of Neural Stem Cell Transgenesis with an Integration-Coupled Gene Expression Switch”. In: *Neuron* 107.4 (Aug. 2020). Publisher: Cell Press, 617–630.e6.

[35] Joerg Schnitzbauer et al. “Super-resolution microscopy with DNA-PAINT”. In: *Nat Protoc* 12.6 (June 2017), pp. 1198–1228.

[36] Minji Jo et al. “Epidermal growth factor receptor-dependent and -independent cell-signaling pathways originating from the urokinase receptor”. In: *J Biol Chem* 278.3 (Jan. 2003), pp. 1642–1646.

[37] Brenda Andrews et al. “Imaging cell biology”. In: *Nat. Cell Biol.* 24.8 (Aug. 2022), pp. 1180–1185.

[38] Baohui Chen et al. “Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system”. In: *Cell* 155.7 (Dec. 2013), pp. 1479–1491.

[39] Hanhui Ma et al. “Multicolor CRISPR labeling of chromosomal loci in human cells”. In: *Proc. Natl. Acad. Sci. U. S. A.* 112.10 (Mar. 2015), pp. 3002–3007.

[40] Hanhui Ma et al. “Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow”. In: *Nat. Biotechnol.* 34.5 (May 2016), pp. 528–530.

[41] Haifeng Wang et al. “CRISPR-mediated live imaging of genome editing and transcription”. In: *Science* 365.6459 (Sept. 2019), pp. 1301–1305.

[42] Julian Grünewald et al. “Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors”. In: *Nature* 569.7756 (May 2019), pp. 433–437.

[43] Martina Fiumara et al. “Genotoxic effects of base and prime editing in human hematopoietic stem cells”. In: *Nat. Biotechnol.* 42.6 (June 2024), pp. 877–891.

[44] Holly A Rees et al. “Analysis and minimization of cellular RNA editing by DNA adenine base editors”. In: *Sci Adv* 5.5 (May 2019), eaax5717.

[45] Jianan Li et al. “Structure-guided engineering of adenine base editor with minimized RNA off-targeting activity”. In: *Nat. Commun.* 12.1 (Apr. 2021), p. 2287.

[46] Changyang Zhou et al. “Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis”. In: *Nature* 571.7764 (July 2019), pp. 275–278.

[47] Monica E Neugebauer et al. “Evolution of an adenine base editor into a small, efficient cytosine base editor with low off-target activity”. In: *Nat. Biotechnol.* 41.5 (May 2023), pp. 673–685.

[48] Livia Zhou et al. “Advances in AAV-mediated gene replacement therapy for pediatric monogenic neurological disorders”. In: *Mol Ther Methods Clin Dev* 32.4 (Dec. 2024), p. 101357.

[49] Jingjing Gao et al. “Gene therapy for CNS disorders: modalities, delivery and translational challenges”. In: *Nat Rev Neurosci* 25.8 (Aug. 2024), pp. 553–572.

[50] Miguel R Chuapoco et al. “Adeno-associated viral vectors for functional intravenous gene transfer throughout the non-human primate brain”. In: *Nat Nanotechnol* 18.10 (Oct. 2023), pp. 1241–1251.

[51] David Goertsen et al. “AAV capsid variants with brain-wide transgene expression and decreased liver targeting after intravenous delivery in mouse and marmoset”. In: *Nat Neurosci* 25.1 (Jan. 2022), pp. 106–115.

[52] Qin Huang et al. “An AAV capsid reprogrammed to bind human transferrin receptor mediates brain-wide gene delivery”. In: *Science* 384.6701 (June 2024), pp. 1220–1227.

[53] X B Liu and E G Jones. “Localization of alpha type II calcium calmodulin-dependent protein kinase at glutamatergic but not gamma-aminobutyric acid (GABAergic) synapses in thalamus and cerebral cortex”. In: *Proc Natl Acad Sci U S A* 93.14 (July 1996), pp. 7332–7336.

[54] A Sík et al. “The absence of a major Ca2+ signaling pathway in GABAergic neurons of the hippocampus”. In: *Proc Natl Acad Sci U S A* 95.6 (Mar. 1998), pp. 3245–3250.

[55] Mojgan Rastegar et al. “MECP2 isoform-specific vectors with regulated expression for Rett syndrome gene therapy”. In: *PLoS One* 4.8 (Aug. 2009), e6810.

[56] Jordane Dimidschstein et al. “A viral strategy for targeting and manipulating interneurons across vertebrate species”. In: *Nat Neurosci* 19.12 (Dec. 2016), pp. 1743–1749.

[57] Jounhong Ryan Cho et al. “Dorsal Raphe Dopamine Neurons Signal Motivational Salience Dependent on Internal State, Expectation, and Behavioral Context”. In: *J Neurosci* 41.12 (Mar. 2021), pp. 2645–2655.

[58] Joseph Cursons et al. “Combinatorial Targeting by MicroRNAs Co-ordinates Post-transcriptional Control of EMT”. In: *Cell Syst* 7.1 (July 2018), 77–91.e7.

[59] Chung-Jung Li et al. “MicroRNA governs bistable cell differentiation and lineage segregation via a noncanonical feedback”. In: *Mol. Syst. Biol.* 17.4 (Apr. 2021), e9945.

[60] Jonathan M Raser and Erin K O’Shea. “Noise in gene expression: origins, consequences, and control”. In: *Science* 309.5743 (Sept. 2005), pp. 2010–2013.

[61] Juan M Pedraza and Alexander van Oudenaarden. “Noise propagation in gene networks”. In: *Science* 307.5717 (Mar. 2005), pp. 1965–1969.

[62] Meaghan Van Alstyne et al. “Gain of toxic function by long-term AAV9-mediated SMN overexpression in the sensorimotor circuit”. In: *Nat. Neurosci.* 24.7 (July 2021), pp. 930–940.

[63] Christian Hinderer et al. “Severe Toxicity in Nonhuman Primates and Piglets Following High-Dose Intravenous Administration of an Adeno-Associated Virus Vector Expressing Human SMN”. In: *Hum. Gene Ther.* 29.3 (Mar. 2018), pp. 285–298.

[64] Stephen E P Smith et al. “Increased gene dosage of Ube3a results in autism traits and decreased glutamate synaptic transmission in mice”. In: *Sci. Transl. Med.* 3.103 (Oct. 2011), 103ra97.

[65] Ann L Collins et al. “Mild overexpression of MeCP2 causes a progressive neurological disorder in mice”. In: *Hum. Mol. Genet.* 13.21 (Nov. 2004), pp. 2679–2689.

[66] Valerie Matagne et al. “Severe offtarget effects following intravenous delivery of AAV9-MECP2 in a female mouse model of Rett syndrome”. In: *Neurobiol. Dis.* 149 (Feb. 2021), p. 105235.

[67] Caleb J Bashor et al. “Engineering the next generation of cell-based therapeutics”. In: *Nat. Rev. Drug Discov.* 21.9 (Sept. 2022), pp. 655–675.

[68] Yuin-Han Loh et al. “Reprogramming of T cells from human peripheral blood”. In: *Cell Stem Cell* 7.1 (July 2010), pp. 15–19.

[69] Haofei Wang et al. “Direct cell reprogramming: approaches, mechanisms and progress”. In: *Nat. Rev. Mol. Cell Biol.* 22.6 (June 2021), pp. 410–424.

[70] Zhongtian Liu et al. “Expression of TNRC6 (GW182) Proteins Is Not Necessary for Gene Silencing by Fully Complementary RNA Duplexes”. In: *Nucleic Acid Ther.* 29.6 (Dec. 2019), pp. 323–334.

[71] Audrius Kilikevicius, Gunter Meister, and David R Corey. “Reexamining assumptions about miRNA-guided gene silencing”. In: *Nucleic Acids Res* 50.2 (Jan. 2022), pp. 617–634.

[72] Hiro-oki Iwakawa and Yukihide Tomari. “Life of RISC: Formation, action, and degradation of RNA-induced silencing complex”. In: *Molecular Cell* 82.1 (2022), pp. 30–43. DOI: 10.1016/j.molcel.2021.11.024.

[73] Pankaj Mehta, Sidhartha Goyal, and Ned S Wingreen. “A quantitative comparison of sRNA-based and protein-based gene regulation”. In: *Molecular Systems Biology* 4.1 (2008), p. 221. DOI: 10.1038/msb.2008.58.

[74] Shankar Mukherji et al. “MicroRNAs can generate thresholds in target gene expression”. In: *Nature Genetics* 43.9 (2011), pp. 854–859. DOI: 10.1038/ng.905.

[75] Dimpal Nyayanit and Chetan J Gadgil. “Mathematical modeling of combinatorial regulation suggests that apparent positive regulation of targets by miRNA could be an artifact resulting from competition for mRNA”. In: *RNA* 21.3 (2015), pp. 307–319. DOI: 10.1261/rna.048652.114.

[76] Ana Kozomara, Maria Birgaoanu, and Sam Griffiths-Jones. “miRBase: from microRNA sequences to function”. In: *Nucleic Acids Res.* 47 (D1 Jan. 2019), pp. D155–D162.

[77] Ana Kozomara and Sam Griffiths-Jones. “miRBase: annotating high confidence microRNAs using deep sequencing data”. In: *Nucleic Acids Res.* 42 (Database issue Jan. 2014), pp. D68–73.

[78] Ana Kozomara and Sam Griffiths-Jones. “miRBase: integrating microRNA annotation and deep-sequencing data”. In: *Nucleic Acids Res.* 39 (Database issue Jan. 2011), pp. D152–7.

[79] Sam Griffiths-Jones et al. “miRBase: microRNA sequences, targets and gene nomenclature”. In: *Nucleic Acids Res.* 34 (Database issue Jan. 2006), pp. D140–4.

[80] Sam Griffiths-Jones. “The microRNA Registry”. In: *Nucleic Acids Res.* 32 (Database issue Jan. 2004), pp. D109–11.

[81] Sam Griffiths-Jones et al. “miRBase: tools for microRNA genomics”. In: *Nucleic Acids Res.* 36 (Database issue Jan. 2008), pp. D154–8.

[82] Victor Ambros et al. “A uniform system for microRNA annotation”. In: *RNA* 9.3 (Mar. 2003), pp. 277–279.

[83] Blake C Meyers et al. “Criteria for annotation of plant MicroRNAs”. In: *Plant Cell* 20.12 (Dec. 2008), pp. 3186–3190.

[84] Sebastian Strauss and Ralf Jungmann. “Up to 100-fold speed-up and multiplexing in optimized DNA-PAINT”. In: *Nat Methods* 17.8 (Aug. 2020), pp. 789–791.

[85] Gabriel Sholder et al. “3’Pool-seq: an optimized cost-efficient and scalable method of whole-transcriptome gene expression profiling”. In: *BMC Genomics* 21.1 (Jan. 2020), p. 64.

[86] Yodai Takei et al. “Multiplexed Dynamic Imaging of Genomic Loci by Combined CRISPR Imaging and DNA Sequential FISH”. In: *Biophys. J.* 112.9 (May 2017), pp. 1773–1776.

[87] Yodai Takei et al. “Integrated spatial genomics reveals global architecture of single nuclei”. In: *Nature* 590.7845 (Feb. 2021), pp. 344–350.

[88] Rosemary C Challis et al. “Systemic AAV vectors for widespread and targeted gene delivery in rodents”. In: *Nat Protoc* 14.2 (Feb. 2019), pp. 379–414.

[89] Johannes Schindelin et al. “Fiji: an open-source platform for biological-image analysis”. In: *Nat. Methods* 9.7 (June 2012), pp. 676–682.

[90] Nicolas L Bray et al. “Near-optimal probabilistic RNA-seq quantification”. In: *Nat. Biotechnol.* 34.5 (May 2016), pp. 525–527.

[91] Brian Munsky, Gregor Neuert, and Alexander van Oudenaarden. “Using gene expression noise to understand gene regulation”. In: *Science* 336.6078 (Apr. 2012), pp. 183–187.

[92] Arjun Raj and Alexander van Oudenaarden. “Single-molecule approaches to stochastic gene expression”. In: *Annu. Rev. Biophys.* 38 (2009), pp. 255–270.

[93] Michael I Love, Wolfgang Huber, and Simon Anders. “Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2”. In: *Genome Biol.* 15.12 (2014), p. 550.

[94] Susanne C M Reinhardt et al. “Ångström-resolution fluorescence microscopy”. In: *Nature* 617.7962 (May 2023), pp. 711–716.

[95] Joschka Hellmeier et al. “Quantification of absolute labeling efficiency at the single-protein level”. In: *Nat Methods* 21.9 (Sept. 2024), pp. 1702–1707.

[96] Pauli Virtanen et al. “Author Correction: SciPy 1.0: fundamental algorithms for scientific computing in Python”. In: *Nat Methods* 17.3 (Mar. 2020), p. 352.

[97] Kendell Clement et al. “CRISPResso2 provides accurate and rapid genome editing sequence analysis”. In: *Nat Biotechnol* 37.3 (Mar. 2019), pp. 224–226.

[98] Anat Kahan et al. “Light-guided sectioning for precise in situ localization and tissue interface analysis for brain-implanted optical fibers and GRIN lenses”. In: *Cell Rep* 36.13 (Sept. 2021), p. 109744.

[99] Carsen Stringer et al. “Cellpose: a generalist algorithm for cellular segmentation”. In: *Nat Methods* 18.1 (Jan. 2021), pp. 100–106.

Chapter 3

MULTICELLULAR CIRCUIT: SYNTHETIC REACTION-DIFFUSION PATTERNING IN MAMMALIAN CELLS

3.1 Summary

During multicellular development, cells form spatially periodic patterns to generate repetitive structures such as digits, vertebrae, and teeth. Turing mechanisms based on short-range self-activation and long-range inhibition have long provided a framework for understanding such processes. However, natural systems involve many interacting signals, making it unclear what minimal circuit architectures can suffice. Here, to address this question, we designed, built, and analyzed modular synthetic reaction-diffusion circuits for self-organized spatial patterning. We compared two circuits composed of both natural (BMP4, Activin) and synthetic (SNIPR) signaling systems with varying diffusion rates, and showed they allow spontaneous patterning on different length scales in mammalian cells. These circuits both implemented classic "local activation with long-range inhibition" regulation and dynamically adapted their active domains for regular patterns. Further, the resulting patterns are tunable via circuit parameters and can adapt to different boundary conditions, consistent with model predictions. Our results demonstrate that simple engineered circuits can recapitulate classical reaction-diffusion (RD) behavior in a cellular context, offering a platform for designing programmable, self-organizing multicellular systems.

3.2 Introduction

Periodic patterns occur throughout biology. They underlie functional structures such as teeth [1], limbs [2] fingerprints [3], and hair [4], and they also enable striking coat patterns such as stripes and spots [5].

In 1952, Alan Turing proposed that periodic biological patterns could arise from local interactions between two diffusible molecules, a short-range activator and a long-range inhibitor, forming what is now known as a reaction-diffusion (RD) system. This framework has since been used to explain natural periodic structures. For instance, Kondo and Asal modeled stripe formation and bifurcation in the marine angelfish *Pomacanthus* using classic RD equations [6]. Similarly, Glover et al. identified a candidate RD mechanism involving BMP and WNT signaling in the

formation of human fingerprints [3].

Despite their generality, RD mechanisms have been difficult to isolate and test directly in vivo. Native patterning circuits often involve many components and are deeply intertwined with developmental programs such as tissue growth and differentiation, making systematic perturbation and analysis challenging. Moreover, not all periodic patterns rely on Turing-like mechanisms, for example, stripes in early *Drosophila* development are driven by transcriptional cascades [7], and zebrafish pigmentation patterns arise from pigment cell migration [8].

A synthetic RD system could address these limitations by providing independence from developmental programs and a tunable platform for studying self-organization. By constructing and analyzing minimal pattern-forming circuits from defined components, we aim to understand the essential features that govern RD driven spatial patterning, including the types of patterns they can generate and the scales over which they operate.

Beyond developmental biology, a synthetic RD system could guide regenerative medicine and its goal of tissue construction and organization [9]. Currently, these efforts focus either on endogenous patterning processes, such as those in organoids [10], or on exogenous, brute-force approaches like 3D printing [11]. Prior efforts have begun to explore synthetic RD circuits in multicellular contexts. For example, Karig et al. achieved patterning in *E. coli* biofilms [12] using two quorum-sensing molecules, while Sekine et al. reconstituted the natural Nodal/Lefty system in mammalian cells [13]. However, these efforts either relied on endogenous molecular competition or were limited in modularity, reducing their utility for engineering diverse or tunable systems.

By providing a scalable alternative for guiding tissue organization, a generalizable, bottom-up strategy for spatial patterning would be immediately useful as well as offering unforeseen benefits. To this end, we designed synthetic RD circuits that generate self-organized spatial patterns in mammalian cells using engineered, orthogonal signaling pathways. In our framework, two signaling molecules, either BMP4 and Activin, or BMP4 and the juxtacrine synthetic Notch receptor (SNIPR), diffuse among cells, while intracellular circuits implement the "local activation and long-range inhibition" principle. Notably, BMP4 serves as either an activator or inhibitor depending on circuit context, demonstrating the flexibility of the design.

These circuits drive the emergence of periodic spatial patterns that are stable yet

reconfigurable. The resulting domains adapt to boundary conditions, shift dynamically under certain parameter regimes, and exhibit length scales corresponding to the diffusion properties of circuit components. Furthermore, pattern features are tunable via circuit parameters, enabling direct control over spatial dynamics. These results establish a modular and versatile platform for engineering synthetic patterning systems and open new opportunities for studying and applying RD principles in developmental and tissue-engineering contexts.

3.3 Results

3.3.1 A framework for modular design of synthetic reaction-diffusion circuits

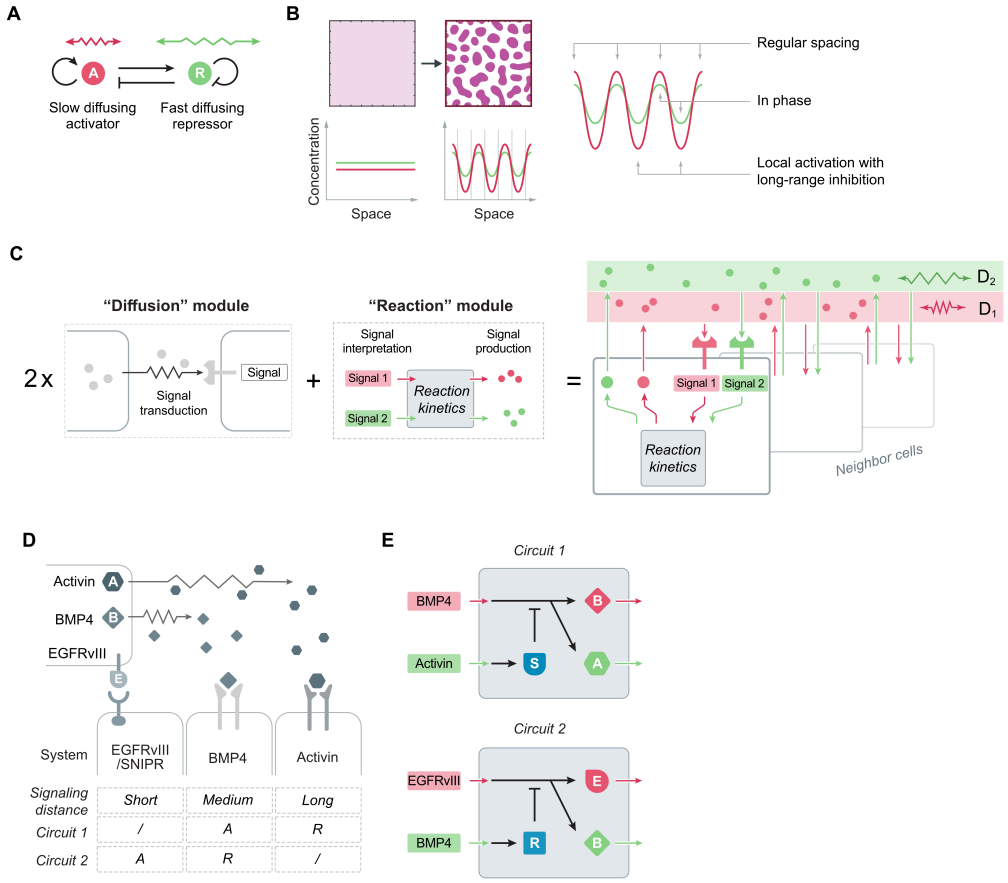

Figure 3.1: Modules for intercellular signal transduction ("diffusion") and intracellular signal integration ("reaction") enable spontaneous spatial patterning.

(A) Classic Turing patterning systems comprise two morphogens: a slower diffusing activator, denoted A, and a more rapidly diffusing repressor, R.

(B) In a kernel-based reaction-diffusion model, these circuits can spontaneously pattern in 1 (lower) and 2 (upper) dimensions.

(C) Turing circuits require diffusion of secretable ligands that can activate cognate responses in other cells (left) and "reaction" modules that implement feedback regulation in which morphogen signals are integrated to control morphogen production or secretion (middle).

(D) Three signaling species with diverse signaling distances are built in two circuits. Circuit 1 uses BMP as the short-range signal and Activin as the long-range signal. Circuit 2 uses SNIPR as the short-range signal and BMP as the long-range signal.

(E) One implementation of reaction kinetics, explored here, involves the activation of both morphogens by the activation morphogens (BMP4 or EGFRvIII, respectively), and the activation of a negative regulator of morphogen production, by the repression morphogens (Activin or BMP4, respectively).

The classical Turing mechanism consists of two components: a slower diffusing species that promotes the production of both components, and a faster diffusing species that inhibits their production (**Figure 3.1A**). These two components diffuse spatially with different rates and interact with one another to generate patterns, based on local activation with long-range inhibition (LALI) [14].

Simulations of such RD circuits, in the absence of diffusion, exhibit a stable, spatially homogenous steady state (**Figure 3.1B**, left). However, this state becomes unstable when diffusion is included, allowing fluctuations to trigger spatial patterning. In this regime, the most rapidly growing spatial mode sets the dominant spatial wavelength of the pattern (**Figure 3.1B**, right). Notably, the resulting patterns are characterized by an "in-phase" relationship between the activator and inhibitor, which peak at the same positions but differ in spatial spread—the inhibitor being more broadly distributed. In this way, by simply tuning the relative length scales of activation and inhibition, a wide range of stable patterns with distinct wavelengths and domain sizes can be realized from the same underlying circuit.

We sought to propose a modular design for such a reaction-diffusion circuit (**Figure 3.1C**). The cells should be able to send and receive two orthogonal signals. Two signal inputs integrate intracellularly and secrete both of them as output.

Many potential molecular systems could implement this type of circuit. Here, we designed two minimal circuits that could be constructed synthetically from three well-characterized, modular pathways, and access a range of patterns across different length scales by varying parameters. In the first circuit, we used BMP4 and Activin—members of the TGF-β superfamily—as morphogens (**Figure 3.1D**, middle and right). These two ligands exhibit distinct diffusive mobilities, and largely orthogonal signaling (**Figure 3.2A, B**). In the second circuit, we used SNIPR, a synthetic Notch-inspired juxtacrine signaling system (**Figure 3.1D**, left), as the activating signal and the same BMP components as the repressor.

In both circuits, the shorter range signal, either BMP4 or SNIPR, acts as the local activator, stimulating its own expression as well as that of the long range inhibitory signal, Activin or BMP4, respectively. The long range inhibitor, in turn, suppresses the activator's production, either by inducing Smad6 expression to inhibit BMP signaling, or by expressing a microRNA targeting EGRRvIII mRNA to repress SNIPR ligand production (**Figure 3.1E**).

Overall, this modular framework represents a new and generalizable strategy for

constructing synthetic RD circuits in mammalian cells.

3.3.2 Synthetic signaling systems exhibit diverse diffusion length scales

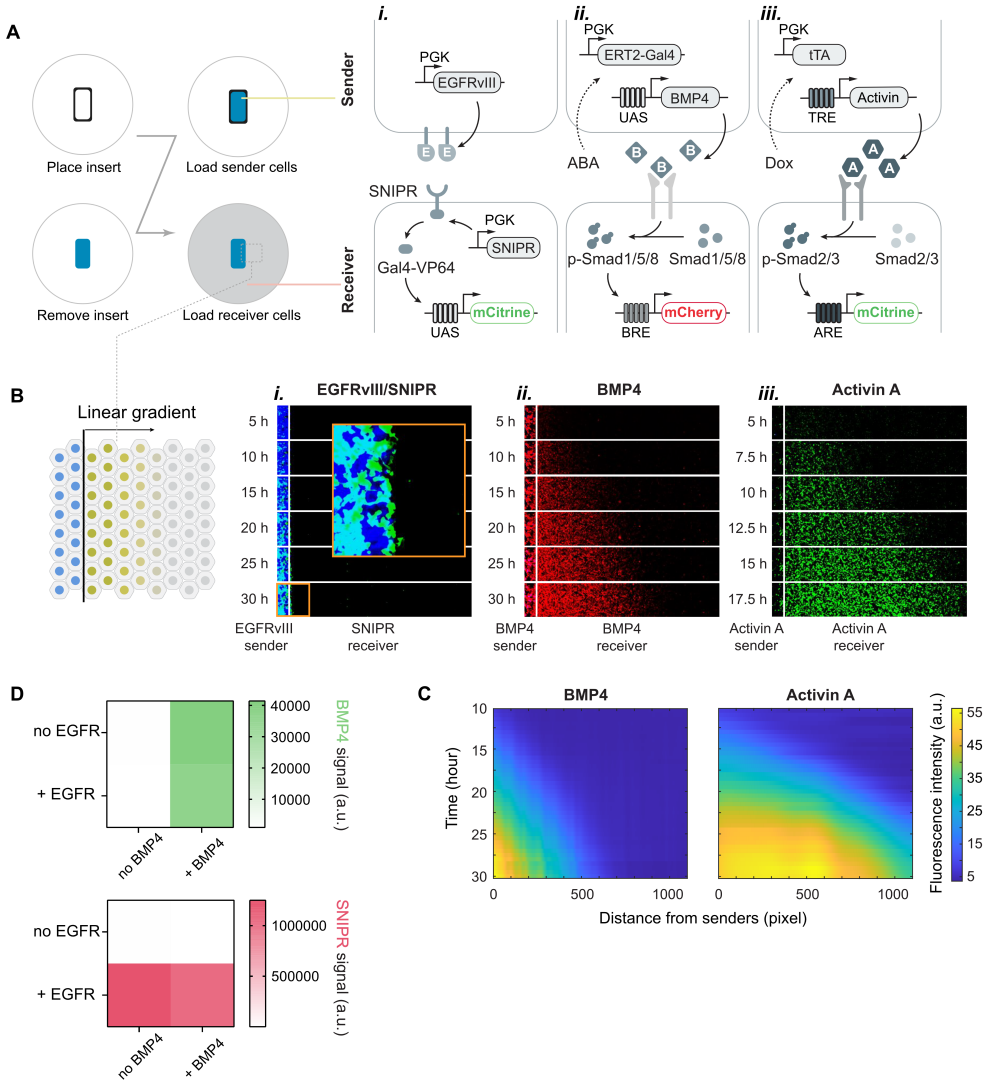

Figure 3.2: Three signaling systems provide a hierarchy of spatial signaling ranges.
(A) The diffusive range can be analyzed using a spatially patterned co-culture sender-receiver system for i) EGFRvIII/SNIPR, ii) BMP4, and iii) Activin A. Sender cell lines (upper row) produce ligands from constitutive (EGFRvIII) or inducible (ABA for BMP4, doxycycline for Activin) promoters. Corresponding receiver cell lines (lower row) ectopically express receptors for the SNIPR pathway or use endogenous receptors for BMP and Activin signaling. They also incorporate ectopic fluorescent protein reporters for each signaling system.
(B) The three systems show different diffusive ranges.
(C) Quantitation of diffusive ranges for each system.

Figure 3.2: (continued)
(D) SNIPR and BMP4 signals are orthogonal. Dual receiver cells containing SNIPR and BMP4 reporter were co-cultured with EGFR sender cells, BMP4 sender cells, and dual sender cells which can send both signals, respectively. Heatmaps indicate the signals of each reporter under labeled experimental conditions.

The three signaling systems exhibit similarities and differences in their signaling mechanisms. BMP4 (**Figure 3.2A**, ii) signals through receptors such as BMPR1A/B and ACTR2/2B, leading to the phosphorylation of Smad1/5/8 effector proteins. Phosphorylated Smad1/5/8 proteins, in turn, co-complex with Smad4 to activate downstream genes containing the BMP response element (BRE) [15, 16, 17]. Activin (**Figure 3.2A**, iii) similarly interacts with a pair of type I receptor ACVR1B and type II receptor ACTR2/2B and phosphorylates a distinct set of Smad proteins, Smad2/3. As with Smad1/5/8, phosphorylated Smad2/3 complexes with Smad4, translocates to the nucleus, binds to Activin response elements (ARE), and activates adjacent genes [18, 19, 20, 21]. By contrast, the third system—SNIPR (**Figure 3.2A**, i)—is a fully synthetic juxtacrine signaling system inspired by the Notch receptor that mediates signaling between adjacent cells (**Figure 3.2A**, left) [22]. Here, signaling cells expressing EGFRvIII on their surface signal to cognate SNIPR receptors whose extracellular domain contains an antibody recognizing EGFRvIII and whose intracellular domain contains the transcriptional activator Gal4-VP64. Upon activation, released Gal4-VP64 translocates to the nucleus, where it binds to and activates UAS promoters.

To directly compare the signaling properties of the three systems, we constructed 'sender' and 'receiver' cell lines for each pathway. We selected AD293 and U2OS cells as platforms because they are equipped with signaling pathway components for natural signal response and can form a firm 2D layer in the culture plate (**Figure 3.2A**). To create BMP and Activin senders we integrated ABA-induced BMP4 and doxycycline-induced Activin genes, respectively. We also constructed a third sender cell line that constitutively expresses EGFRvIII, which acts as an engineered ligand for the cognate SNIPR.

To monitor signaling, we constructed a set of three cognate receiver cell lines, each expressing a fluorescent protein reporter under the control of one pathway. The BMP and Activin receiver lines expressed mCherry or mCitrine under the control of BMP or Activin response elements, respectively (**Figure 3.2A**). Similarly, in U2OS

cells, we built SNIPR receiver cells that activate YFP in response to stimulation by EGFRvIII expressing cells.

Using these cell lines, we quantified spatio-temporal signal transmission among cells. We co-cultured each sender-receiver pair using an iBiDi insert in conjunction with sequential plating to create a sharp sender-receiver interface (**Figure 3.2A**, left), while recording time-lapse movies. In this quasi-1D geometry, the signal perception of the receivers, as revealed by their reporter genes, exhibited temporally increasing spatial signaling gradients (**Figure 3.2B**, left).

Analysis of the resulting signal gradients revealed distinct diffusive properties for BMP4 and Activin, with effective diffusion coefficients of 3.7 $\mu m^2/s$ and 16.3 $\mu m^2/s$, respectively (**Figure 3.2C**). These data are consistent with previous works showing that Activin diffuses faster than BMP4 [23, 24]. By contrast, the EGFRvIII-presenting cells could only activate the SNIPR reporter on directly adjacent cells via juxtacrine signaling. Thus, the three signaling systems exhibited widely varying lengthscales, as anticipated.

To evaluate the independence (orthogonality) of these signaling channels, we engineered AD293 cells with two distinct fluorescent protein reporters, controlled by the BRE or ARE response elements. We titrated and quantified the responses to recombinant BMP4 and Activin ligands. Despite minor inhibition of Activin in the high concentration regime—consistent with previous observations [25, 26]—BMP4 and Activin are generally orthogonal to each other [27]. Additionally, in U2OS cell lines, the EGFR senders efficiently activated the SNIPR reporter but did not interfere with BMP4 signaling, and vice versa (**Figure 3.2D**).

3.3.3 SNIPR enables tunable, cell density-dependent self-activation

To experimentally identify the parameter regimes that allow pattern formation, it is desirable for each signaling component to be independently tunable. Unlike diffusible morphogens such as BMP4 and Activin whose effective concentrations can be readily modulated by adding recombinant ligands or small-molecule inhibitors, SNIPR is a membrane-tethered receptor that signals through direct cell–cell contact. This property precludes simple concentration tuning and motivates alternative strategies to modulate SNIPR activity.

We first explored whether SNIPR signaling could be tuned by varying the ratio of sender to receiver cells. Altering this ratio changes the probability of productive cell–cell contacts, thereby enabling control over both the fraction of activated re-

ceiver cells and their average activation level (**Figure 3.3A**). In addition, we hypothesized that the effective removal rate of SNIPR could modulate signaling strength. To test this, we fused a Shield-1–stabilizable degron derived from FKBP12 [28] to the C-terminus of SNIPR. Consistent with this hypothesis, we observed drug concentration–dependent activation: increasing Shield-1 levels led to both a larger fraction of activated cells and stronger activation signals (**Figure 3.3B**).

Previous studies have shown that the Notch signaling system can propagate signals spatially and temporally through mutual activation, both in the native Notch–Delta context [29] and in synthetic Notch-based systems [30]. In these implementations, however, signal propagation is initiated by a constitutive source of sender cells. Whether an initially homogeneous population can spontaneously activate itself remains unresolved. To address this question, we engineered a monoclonal cell line that constitutively expresses SNIPR, with the SNIPR intracellular domain driving expression of its own ligand, EGFRvIII (**Figure 3.3C**, left). We reasoned that such a positive-feedback circuit could enable an initially "OFF" population to self-activate over time (**Figure 3.3C**, right).

Indeed, this monoclonal population exhibited pronounced density-dependent self-activation (**Figure 3.3D**): higher initial seeding densities resulted in a larger fraction of activated cells. Notably, the circuit displayed a form of hysteresis, in which cells tended to remain in their original state. Populations seeded at low density remained largely inactive even after reaching confluence by the end of the experiment, whereas populations seeded at high density robustly activated and maintained the ON state.

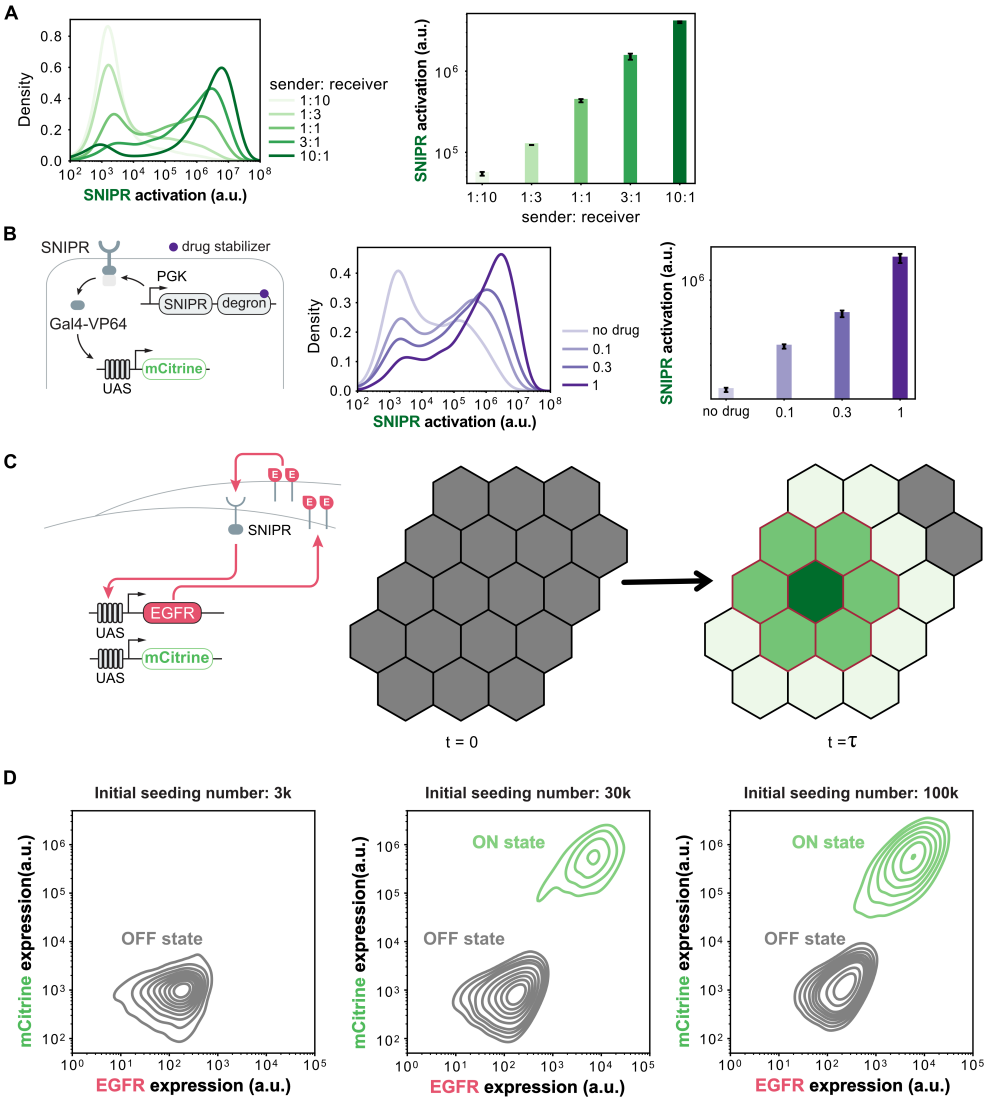

Figure 3.3: SNIPR enables tunable, cell density-dependent self-activation.
(A) Sender-to-receiver ratios tune SNIPR activation. Left, probability density function of the SNIPR reporters' fluorescence intensities. Right, median fluorescence intensities of the SNIPR reporter in the receiver line. Data represents three biological replicates. Error bars, geometric standard deviation. Sender cells were gated out by another fluorescence protein indicator.
(B) Small molecule drug stabilizes SNIPR and leads to a dose-dependent activation. Left, schematics of the drug-titratable, degron-tethered SNIPR. Middle, probability density function of the SNIPR reporters' fluorescence intensities. Right, median fluorescence intensities of the SNIPR reporter in the receiver line. Data represents three biological replicates. Error bars, geometric standard deviation. Sender cells were gated out by another fluorescence protein indicator. The drug is AquaShield-1, and the unit of drug concentration is μM.

Figure 3.3: (continued)
(C) Design of the self-activatable SNIPR circuit. Red edges on the right represent EGFR expression on the membrane. Green color gradients show different activation levels of the SNIPR reporter.
(D) SNIPR circuit enables initial seeding density-dependent activation. Contour plots show the EGFR (x-axis) and the reporter (y-axis) expressions. Contours beyond 25% are plotted.

3.3.4 Synthetic reaction-diffusion circuits enable diverse configurations.

Creating the Turing circuits requires (1) choosing two morphogens with different diffusion coefficients, (2) allowing the more slowly diffusing species to activate expression of both morphogens, and (3) allowing the more rapidly diffusing species to activate a repressor of the activator (**Figure 3.1E**). For the BMP4/Activin circuit, we engineered the slower diffusing BMP4 to activate expression of both morphogens, and the faster diffusing Activin to activate expression of Smad6, which is an inhibitory Smad that down-regulates BMP4, but not activin, signaling (**Figure 3.4A**). To create the SNIPR/BMP4 circuit, we took a similar approach, with BMP4 now playing the role of the faster diffusing species, repressing the SNIPR ligand by inducing a synthetic miRNA [31], which targets a specific engineered binding sequence in the EGFRvIII ligand gene to downregulate ligand production (**Figure 3.4B**). This miRNA-based strategy is generalizable and could be used to wire reaction diffusion kinetics among diverse signaling pathways.

To track circuit dynamics, we designed a post-translationally regulated reporter that allows modulation of reporter half-life with the drug trimethoprim. The fluorescent reporter gene was integrated within the same transcript as its corresponding patterning component, using an internal ribosome entry sites (IRES) sequence for independent expression. Additionally, we fused the H2B protein to an N-terminal trimethoprim (TMP)-inducible DHFR protein degron. Without TMP, the reporter is rapidly degraded, but with TMP it is stabilized and can accumulate over time (**Figure 3.4A, B**). This design ensures cells only accumulate a detectable level of fluorescent signal when TMP is added and prevents accumulation of fluorescent protein prior to the experiment. For the first circuit, we used mCherry and mTurquoise to label the production of short and long-range signals, respectively, while mCitrine was used to track the third intracellular inhibitory component.

The first circuit must satisfy two conditions for patterning: First, BMP4 signaling should be able to access high enough levels to sustain high BMP4 production.

Second, Activin-induced Smad6 production should be able to block BMP4 self-activation. To satisfy these conditions, we selected stable clones of wild-type AD293 cells harboring the self-activated BMP4 and Activin-induced Smad6 expression constructs (**Figure 3.4C**). Stable cell lines with these integrations strongly activated BMP4 without requiring recombinant ligand, indicating that BMP4 can self-activate. Addition of recombinant Activin was nevertheless able to inhibit BMP4 self-activation in a dose-dependent manner (**Figure 3.4C**). These "open loop" results demonstrate the required features of self-activation and inhibition by BMP4 and Activin, respectively.

Similar tests with the SNIPR system prove the positive-regulating circuit with SNIPR and its ligand EGFRvIII can also activate the cells in the culture condition. BMP4 induces transcription of miRNA, which specifically represses the mRNA for EGFRvIII ligand production. To quantify the EGFRvIII expression level titrated by BMP4, we used siRNA to target the BMP4 transcript in the full circuit monoclonal line, and applied recombinant BMP4 into the culture media. With no or low concentrations of recombinant BMP4, the circuit maintains high EGFR expression; when BMP4 keeps increasing, the amount of miRNA expression is elevated, and gradually turns off the EGFR expression (**Figure 3.4D**).

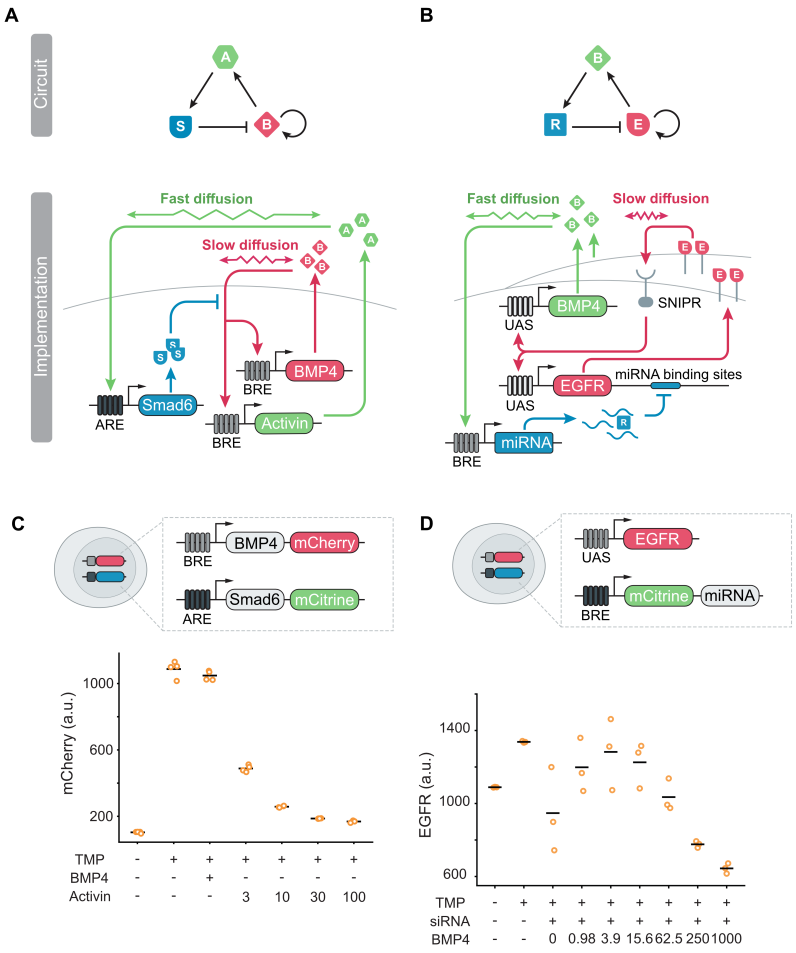

Figure 3.4: Synthetic reaction-diffusion circuits enable diverse configurations.
(A) The BMP4/Activin circuit design includes BMP4 for short-range signaling, Activin for long-range signaling, and Smad6 as the intracellular inhibitive component.
(B) The SNIPR/BMP4 circuit design includes EGFRvIII/SNIPR for short-range signaling, BMP4 for long-range signaling, and microRNA as the intracellular inhibitive component.
(C) Evaluation with an open-loop circuit monoclonal validates the circuit design. The BMP4 self-regulation module can maintain cells at the activation state. Activin-induced Smad6 can further repress the BMP self-activation. Data represents three biological replicates.
(D) Evaluation with a full circuit monoclonal validates the circuit design. EGFR expression is measured by EGFR antibody staining. The SNIPR-EGFR self-activation module can maintain cells at the activation state with siRNA targeting the BMP4 transcript. Escalation of recombinant BMP4 concentration in the media induces miRNA expression and gradually tunes down EGFR expression. Data represents three biological replicates.

3.3.5 Synthetic circuits produce spatial patterns in cell culture

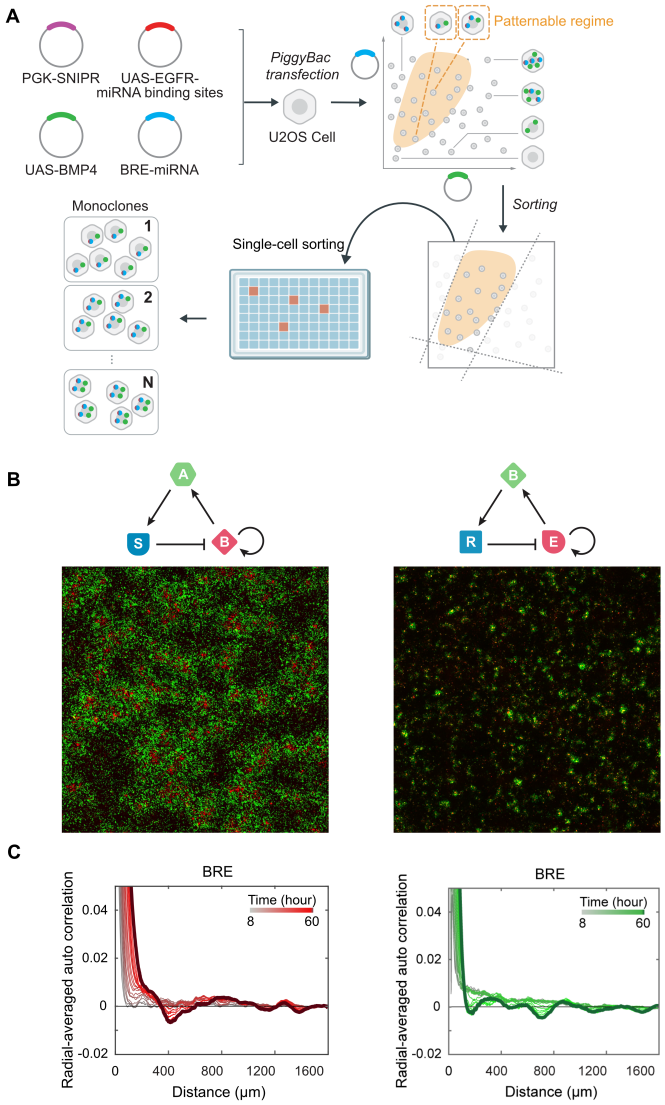

Figure 3.5: Both synthetic circuits generate spatial patterns.
(A) Workflow for constructing SNIPR-EGFR patterning cell lines. A PiggyBac-based transfection of the indicated plasmids generates heterogeneous integrations. Fluorescence-activated cell sorting (FACS) isolates clones whose expression ranges are conducive to pattern formation (yellow region).
(B) Under patterning conditions, the monoclonal cells in both circuits develop non-periodic spatial patterns. Left, the BMP-Activin circuit exhibits localized regions of high BMP signaling (red areas) separated by regions of low BMP signaling, so as the Activin signaling (green areas). Right, the SNIPR-BMP circuit displays regions of high SNIPR signaling (red areas) separated by regions of low SNIPR signaling, so as the BMP signaling (green areas). Scale bar, 500 μm.
(C) Autocorrelation analysis of both circuits (left: BMP-Activin; right: SNIPR-BMP) confirms the spatial regularity of the final pattern, with a characteristic domain spacing emerging as the pattern stabilizes.

The open-loop circuits revealed that self-activation and long range inhibition were both functional. We therefore proceeded to build the full circuit and identify clones in the parameter regime required for effective patterning. To do so, we used a "shotgun" strategy, inserting circuit genes randomly using the piggyBac transposon and identifying monoclones that were capable of patterning (**Figure 3.5A, Methods**).

With clones in hand, we performed pattern formation assays. Briefly, before patterning, we individually cultured each monoclonal line to avoid pre-patterning prior to the experiment. More specifically, for the first circuit, we added the BMP inhibitor Noggin to maintain cells in the "off" state, while for the second circuit, we maintained the SNIPR-FKBDD culture into the Shield-1-free media. Then we replated these cells and allowed them to settle with Noggin protection. After cells attached to the plate surface, we proceeded to the pattern formation assay. These procedures ensure that all cells begin patterning with identical genotype, no morphogen production, and in a confluent state, with limited cell divisions during patterning.

After washing away Noggin, we initiated the experiment in a pattern-formation medium containing TMP (for fluorescent protein stabilization), Palbociclib (to inhibit cell division), and specific concentrations of small-molecule drugs to tune circuit parameters. As shown in **Figure 3.5B**, our BMP-Activin circuit produces a typical pattern: after 60 hours, cells that were initially homogeneous exhibit spatially heterogeneous activation in both the mCherry (BRE promoter) and mCitrine (ARE promoter) channels. Our SNIPR-BMP circuit also produces a typical pattern with a smaller lengthscale compared to BMP-Activin circuit, in both the mCherry (UAS promoter) and mCitrine (BRE promoter) channels.

Finally, we computed the autocorrelation of the patterns, which measures how each image correlates with itself as a function of shift (**Figure 3.5C**). After radially averaging to remove directional biases, both circuits' autocorrelation curves converged to a stable shape, corresponding to an average cluster size of about 400 μm for the BMP-Activin circuit, and about 180 μm for the SNIPR-BMP circuit, which is consistent with the classic Turing model that the wavelengths increase with diffusion coefficients [32]. Notably, a sharper trough at short distances in the early mCherry autocorrelation disappeared at later stages, reflecting competition and spacing adjustments among active domains.

Together, these analyses confirm that the synthetic reaction-diffusion circuit can drive pattern formation through local activation with long-range inhibition.

3.4 Discussion

Self organized spatial patterning is a fundamental feature of multicellular development. However, isolating and testing the minimal principles sufficient for generating such patterns has been challenging. Here, we engineered and analyzed synthetic reaction diffusion (RD) circuits in mammalian cells that produce stable, self organized spatial patterns using defined molecular components. The core design principle is modularity: discrete units that can be independently designed, implemented, and recombined to perform distinct functions.

With the deployment of two synthetic circuits that facilitate spatial patterning in cell cultures, we have substantiated a universal modular design for synthetic reaction-diffusion circuits applicable to multicellular mammalian systems. Utilizing a transient reporter system that accurately captures the spatial-temporal dynamics, our approach elucidates the intricate details of self-organization. Key findings include: Each individual pattern exemplifies the principle of "local activation with long-range inhibition," demonstrating dynamic adaptation in their positioning and interaction within the system. The modulation of circuit parameters and adaptation to varying boundary conditions affirm the tunability of these patterns, which is rigorously validated through perturbation studies. Comparative analysis between the two circuits highlights distinct spatial scales of patterning, which correlate directly with the diffusion capabilities of their respective circuit components. Within specific parameter regimes, dynamic patterns emerge consistently across both circuits, illustrating the robustness and versatility of our synthetic designs. These insights not only validate the effectiveness of our synthetic circuits but also pave the way for broader applications in understanding and manipulating biological systems through engineered synthetic modules.

Our RD system design is highly extensible. In the two circuit architectures explored here, BMP4 functions either as a slow-diffusing activator or as a fast-diffusing repressor, illustrating a general principle: RD pattern formation is determined primarily by the regulatory topology of the circuit rather than by the specific molecular identity of the morphogen. Under this framework, the intercellular signaling components used in our system could, in principle, be replaced by other natural or synthetic signaling elements, provided that they exhibit sufficiently distinct diffusion properties. Likewise, the intracellular reaction modules are modular and amenable to further engineering. Beyond transcriptional regulation, alternative regulatory layers, including dimerization-based transcriptional circuits [33] and protease-based

post-translational circuits [34]—could be incorporated to expand the design space and dynamic repertoire of RD behaviors.

In the broader context of synthetic tissue engineering, our patterning circuit provides a general strategy for breaking symmetry in initially homogeneous multicellular assemblies. By inducing spatially correlated bifurcation of cell states, the circuit enables spatially coordinated fractional control, a key requirement for engineering functional tissues composed of multiple cell types. For instance, coupling our patterning module with appropriate functional differentiation programs could, in principle, allow the generation of a prescribed number and spatial distribution of islets of Langerhans in a synthetic pancreas. More broadly, integration of the RD circuit with active mechanical processes—such as cell migration, force generation, or matrix remodeling—could further extend its utility to patterning higher-order structures, including vascular networks within synthetic organs.

Despite these capabilities, our system also has several limitations. First, in a closed culture environment such as a petri dish, signaling molecules inevitably accumulate and saturate over time, which hinders long-term stabilization of the final pattern. Second, the implementation of multi-component synthetic circuits imposes a metabolic burden on cells, potentially constraining robustness and scalability. These limitations point to clear directions for future work. In particular, integration with microfluidic platforms could provide continuous nutrient supply and waste removal, helping maintain metabolic homeostasis while preventing uncontrolled signal accumulation. Importantly, such systems could be designed to preserve the positional information established by morphogen gradients, thereby enabling sustained and well-controlled pattern formation in extended culture.

Together, our work introduced a modular design approach to mammalian multicellular reaction-diffusion patterning and showcased the dynamic pattern formation capabilities of two distinct synthetic circuits. These findings not only confirm the functionality and effectiveness of our synthetic designs but also broaden the potential for utilizing engineered synthetic modules to deepen our understanding and enhance our ability to manipulate biological systems.

3.5 Methods

3.5.1 Simulation

The simulation in this work (presented in Figure 3.1B) is based on the kernel-based reaction-diffusion patterning simulation [35], which is a non-partial differentiation

equation (PDE)-represented model that can generate all standard variations of stable 2D patterns.

3.5.2 Molecular cloning

Constructs used in this study were generated using standard cloning procedures. The inserts were generated using PCR or gBlock synthesis (IDT) and were ligated either by T4 ligase (NEB #M0202M) or In-Fusion (Takara #102518) or Gibson assembly with backbones linearized using restriction enzyme digestion.

3.5.3 Cell culture

AD293 (Agilent #240085), a derivative of the standard HEK293 (Human Embryonic Kidney) cell line, was chosen for its improved cell adherence and plaque formation characteristics in the first circuit. Cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS), 1×Penicillin/Streptomycin/L-glutamine, 1 mM Sodium Pyruvate, and 1 mM Non-Essential Amino Acids (NEAA). Cells were passaged every three days, with a ratio of 1:10. For the maintenance of cells containing the BMP4 self-activation circuit, an additional 250 ng/mL of recombinant Noggin was included in the media to ensure the circuit remained off.

U2OS cells (ATCC #HTB-96) were chosen for its contact inhibition property and suitability for microscopy imaging in the second circuit. Cells were cultured in DMEM (Dulbecco's Modified Eagle Medium, ThermoFisher #11960-069) supplemented with 10% FBS, 1 U/mL penicillin, 1 µg/mL streptomycin, 1 mM sodium pyruvate, 1×NEAA (ThermoFisher #11140-050), 1 mM L-glutamine, and 0.1 mg/mL Normocin (InvivoGen #ant-nr). For the sender-receiver co-culture experiment, sender cells and receiver cells were seeded at the labeled ratio, and co-cultured for 72 hrs post-seeding.

When imaging was required, the regular DMEM was replaced with FluoroBrite™ DMEM from Gibco, and 0.1 µM TMP and 5 µM Palbociclib were added.

3.5.4 Cell line construction

Cell lines were constructed by using the PiggyBac transposase-based method. Cells were co-transfected with the PiggyBac-backbone constructs and the transposase construct, either using Lipofectamine LTX (ThermoFisher) for AD293 or Fugene HD (Promega #E2311) for U2OS, according to the manufacturer's protocol. Antibiotics (Hygromycin, Puromycin, and Blasticidin) were added the following day

post-transfection for selection according to the antibiotics markers the constructs carried. Antibiotics-selected polyclonals were then re-plated into the experimental media for sorting.

3.5.5 Flow cytometry

Cells were incubated 2 days or 3 days, depending on the experimental settings, before the flow cytometry experiments. Cells were trypsinized with 50 µL of either 0.05% trypsin (for AD293 cells) or 0.25% trypsin (for U2OS cells) for 5 minutes at 37°C. After digestion, cells were resuspended with 150 µL of HBSS containing 2.5 mg/mL BSA and 1 mM EDTA. Cells were then filtered through a 40 µm cell strainer and analyzed using a CytoFLEX S instrument (Beckman Coulter). We used the FlowJo V10 and self-build python code to analyze the flow data.

3.5.6 Cell sorting

To prepare the stable cell lines, cells were harvested and resuspended in the sorting buffer (BD FACS Pre-Sort Buffer) supplemented with 1 U/mL DNAse I by the cell sorter (Sony MA900) as mono-clones using 96-well plates, or poly-clones using the 1.5 mL eppendorf tubes. Cells were expanded in the 24 well plate and validated phenotypes before further measurements.

3.5.7 Reconstitution of morphogen gradients

Poly-D-lysine (0.1mg/mL) was mixed with DPBS at 1:1 ratio. 1 mL of the mixture was loaded in each well of a 6-well plate and incubated at room temperature for 1 hour. The coating solution was then removed, and the plate was rinsed three times with 1.5 mL of ddH2O and dried, sealed, and stocked in 4C for up to two weeks for use. An iBiDi insert (#80209) was carefully positioned and gently pressed at the center of the well.

Senders were trypsinized, resuspended, and seeded at 80k per 100 µL media for each iBiDi insert. After 12 to 24 hours post-seeding, the iBiDi insert was carefully removed by a tweezer and rinsed once with 1 mL media. Receivers were then loaded at 0.6M per 1 mL media for each well. After receivers settled down, media was removed and the well was rinsed once with 1 mL imaging media. A mix of 1.5 mL imaging media encompassing small-molecule inducers and 0.5 mL Matrigel were loaded on cells. The plate was settled in the incubator for 15 min and transferred to the EVOS incubator afterwards for time-lapse movies.

48 hrs of time-lapse movies were taken and the samples were scanned every 30

min. For each well, 10×5 field of views (FOVs) were scanned, with phase channel adjusting the autofocus, BFP channel locating the sender-receiver boundary, and YFP or mCherry signal indicating the signal gradients.

3.5.8 Patterning experiment setup

100 µL 50 mg/mL D-poly-lysine solution was added to a well of a 96-well plate for 1 hr at room temperature. Afterwards, the plate was thoroughly washed with 200 µL ddH_2O three times to remove any residual D-poly-lysine, and air-dried at room temperature for 24 hrs. For the first circuit, cells were cultured in 150 µL media with 250 ng/mL Noggin for 24 hrs and rinsed with 300 µL regular media before seeding; for the second circuit, cells were cultured in a AquaShield-1-free media for at least 24 hrs before seeding, both to prevent pre-patterning. Patterning monoclonals were then seeded at full confluence into the prepared plates in 500 µL patterning media. Cells were placed under an environmental chamber-equipped Olympus microscope for time-lapse movie capture.

Time-lapse movies were taken by an inverted Olympus IX81 fluorescence microscope with an ASI 2000YX automated stage, an iKon-M charge-coupled device camera (Andor), a 20× dry objective lens (0.7 NA), and an X-Cite XLED1 light source (Lumen Dynamics). Microscope and image acquisition were controlled by Metamorph software (Molecular Devices). Cells were kept in a customized environmental chamber with a humidified, 37°C and 5% CO2 atmosphere. For each well, a square area at the well center was imaged. The region consists of 25 FOVs (5×5), and each FOV is 650 µm × 650 µm, with a 50-µm overlap between two adjacent FOVs. Movies were started 1 hour after induction and taken every 1 hour.

References

[1] Kathryn D Kavanagh, Alistair R Evans, and Jukka Jernvall. “Predicting evolutionary patterns of mammalian teeth from development”. In: *Nature* 449.7161 (Sept. 2007), pp. 427–432.

[2] Kimberly L Cooper. “Self-organization in the limb: a Turing mechanism for digit development”. In: *Curr Opin Genet Dev* 32 (June 2015), pp. 92–97.

[3] James D Glover et al. “The developmental basis of fingerprint pattern formation and variation”. In: *Cell* 186.5 (Mar. 2023), 940–956.e20.

[4] H S Jung et al. “Local inhibitory action of BMPs and their relationships with activators in feather formation: implications for periodic patterning”. In: *Dev Biol* 196.1 (Apr. 1998), pp. 11–23.

[5] Shigeru Kondo, Motoko Iwashita, and Motoomi Yamaguchi. “How animals get their skin patterns: fish pigment pattern as a live Turing wave”. In: *Int J Dev Biol* 53.5 (2009), pp. 851–856.

[6] S Kondo and R Asal. “A reaction-diffusion wave on the skin of the marine angelfish Pomacanthus”. In: *Nature* 376.6543 (Aug. 1995), pp. 765–768.

[7] M J Pankratz and H Jäckle. “Making stripes in the Drosophila embryo”. In: *Trends Genet* 6.9 (Sept. 1990), pp. 287–292.

[8] Masakatsu Watanabe and Shigeru Kondo. “Is pigment patterning in fish skin determined by the Turing mechanism?” In: *Trends Genet* 31.2 (Feb. 2015), pp. 88–96.

[9] Ricard Solé et al. “Open problems in synthetic multicellularity”. In: *NPJ Syst Biol Appl* 10.1 (Dec. 2024), p. 151.

[10] Zixuan Zhao et al. “Organoids”. In: *Nat Rev Methods Primers* 2 (Dec. 2022).

[11] Madhuri Dey and Ibrahim T Ozbolat. “3D bioprinting of cells, tissues and organs”. In: *Sci Rep* 10.1 (Aug. 2020), p. 14023.

[12] David Karig et al. “Stochastic Turing patterns in a synthetic bacterial population”. In: *Proc Natl Acad Sci U S A* 115.26 (June 2018), pp. 6572–6577.

[13] Ryoji Sekine, Tatsuo Shibata, and Miki Ebisuya. “Synthetic mammalian pattern formation driven by differential diffusivity of Nodal and Lefty”. In: *Nat Commun* 9.1 (Dec. 2018), p. 5456.

[14] A Gierer and H Meinhardt. “A theory of biological pattern formation”. In: *Kybernetik* 12.1 (Dec. 1972), pp. 30–39.

[15] C H Heldin, K Miyazono, and P ten Dijke. “TGF-beta signalling from cell membrane to nucleus through SMAD proteins”. In: *Nature* 390.6659 (Dec. 1997), pp. 465–471.

[16] J Massagué. “TGF-beta signal transduction”. In: *Annu Rev Biochem* 67 (1998), pp. 753–791.

[17] Olexander Korchynskyi and Peter ten Dijke. “Identification and functional characterization of distinct critically important bone morphogenetic protein-specific response elements in the Id1 promoter”. In: *J Biol Chem* 277.7 (Feb. 2002), pp. 4883–4891.

[18] Di Chen, Ming Zhao, and Gregory R Mundy. “Bone morphogenetic proteins”. In: *Growth Factors* 22.4 (Dec. 2004), pp. 233–241.

[19] M Kawabata and K Miyazono. “Signal transduction of the TGF-beta superfamily by Smad proteins”. In: *J Biochem* 125.1 (Jan. 1999), pp. 9–16.

[20] K Miyazono, P ten Dijke, and C H Heldin. “TGF-beta signaling by Smad proteins”. In: *Adv Immunol* 75 (2000), pp. 115–157.

[21] Joan Massagué. "TGFβ signalling in context". In: *Nat Rev Mol Cell Biol* 13.10 (Oct. 2012), pp. 616–630.

[22] Iowis Zhu et al. "Modular design of synthetic receptors for programmed gene regulation in cell therapies". In: *Cell* 185.8 (Apr. 2022), 1431–1443.e16.

[23] N McDowell, J B Gurdon, and D J Grainger. "Formation of a functional morphogen gradient by a passive process in tissue from the early Xenopus embryo". In: *Int J Dev Biol* 45.1 (2001), pp. 199–207.

[24] C M Jones, N Armes, and J C Smith. "Signalling by TGF-beta family members: short-range effects of Xnr-2 and BMP-4 contrast with the long-range effects of activin". In: *Curr Biol* 6.11 (Nov. 1996), pp. 1468–1475.

[25] Oddrun Elise Olsen et al. "Activin A inhibits BMP-signaling by binding ACVR2A and ACVR2B". In: *Cell Commun Signal* 13 (June 2015), p. 27.

[26] Szabina Szófia Szilágyi, Orit Gutman, and Yoav I Henis. "Complex Formation Among TGF-β Receptors in Live Cell Membranes Measured by Patch-FRAP". In: *Methods Mol Biol* 2488 (2022), pp. 23–34.

[27] Sheng Wang. "Synthetic Circuits for Multicellular Spatial Patterning". PhD thesis. California Institute of Technology, 2023.

[28] Laura A Banaszynski et al. "A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules". In: *Cell* 126.5 (Sept. 2006), pp. 995–1004.

[29] Mitsuhiro Matsuda et al. "Synthetic signal propagation through direct cell-cell interaction". In: *Sci Signal* 5.220 (Apr. 2012), ra31.

[30] Marco Santorelli et al. "Control of spatio-temporal patterning via cell growth in a multicellular synthetic gene circuit". In: *Nat Commun* 15.1 (Nov. 2024), p. 9867.

[31] Rongrong Du, Michael J. Flynn, Karan Mahe, Monique Honsa, Bo Gu, Dongyang Li, Sean E. McGeary, Viviana Gradinaru, Ralf Jungmann, and Michael B. Elowitz. "miRNA modules for precise, tunable control of gene expression". In: *Molecular Cell* 86(1) (2026), 194–212.e7. DOI: 10.1016/j.molcel.2025.11.028.

[32] Qi Ouyang et al. "Dependence of Turing pattern wavelength on diffusion rate". In: *J. Chem. Phys.* 102.6 (Feb. 1995). Publisher: AIP Publishing, pp. 2551–2555.

[33] Ronghui Zhu et al. "Synthetic multistability in mammalian cells". In: *Science* 375.6578 (Jan. 2022), eabg9765.

[34] Xiaojing J Gao et al. "Programmable protein circuits in living cells". In: *Science* 361.6408 (Sept. 2018), pp. 1252–1258.

[35] Shigeru Kondo. “An updated kernel-based Turing model for studying the mechanisms of biological pattern formation”. In: *J Theor Biol* 414 (Feb. 2017), pp. 120–127.

Chapter 4

CONCLUDING REMARKS

4.1 Living matters are dynamical systems

In this thesis, I introduced two synthetic circuits operating at distinct length scales. In Chapter 2, I constructed a single-gene regulatory circuit that enables precise and tunable control of gene expression. In Chapter 3, I built a multicellular circuit capable of generating periodic spatial patterns. Although these two systems differ markedly in scale and implementation, they are united by a common conceptual premise: living matter can be productively viewed as a dynamical system—one that can be analyzed, perturbed, and, to some extent, re-designed in a predictive manner.

In both cases, I adopted kinetic rather than thermodynamic models as the primary theoretical framework. Kinetic models are appealing in their simplicity and directness: they specify explicit molecular species and reaction flows that an engineer can, at least in principle, pursue experimentally. However, despite often being regarded as "mechanistic," many commonly used kinetic descriptions are in fact phenomenological. They typically assume the existence, direction, and functional form of reactions without explicitly encoding the energetic or statistical principles that give rise to them.

This limitation is evident in the model of the DIMMER system described in Chapter 2. A simplified kinetic model was used to capture the effective binding and unbinding dynamics between RISC and target mRNAs, together with catalytic repression. While this model successfully explains system-level behavior, it does not fully resolve a key molecular puzzle: why a perfectly complementary target exhibits poorer dosage compensation than multimerized, partially complementary targets in our interested dynamic ranges. We know experimentally, the latter relies on the TNRC6-dependent, multivalent regulation, while the former one does not. Thereby, a more complete explanation likely requires a statistical-mechanical description of the interaction tendencies among target RNAs, RISC, and scaffold proteins. Developing such a model would in turn demand more detailed experimental inputs, such as single-molecule binding kinetics and slicing rates, which are becoming increasingly feasible given advances in *in vitro* reconstitution and single-molecule imaging. Nevertheless, how faithfully these measurements reflect the crowded and

heterogeneous *in vivo* environment remains an open question. Bridging kinetic and thermodynamic perspectives, therefore, represents not merely a modeling challenge, but an ongoing dialogue between theory and experiment.

Across the two systems studied here, my emphasis on dynamics took different forms. In the DIMMER circuit, the primary focus was on steady-state outcomes. In contrast, for the RD circuit, dynamical trajectories were also of interest, even though the ultimate steady-state pattern remained a central object of study—reflecting the fact that most living organisms exhibit stable spatial organization rather than continuously fluctuating patterns. Yet dynamics are indispensable to biological form and function. During development, organisms repeatedly generate new structures and dismantle old ones, navigating transitions that are neither instantaneous nor strictly irreversible.

That said, *metastability* may offer a more faithful description of biological cell states than strict stability. Metastable states allow a cell to maintain its current identity while retaining the capacity to transition into alternative states in a partially autonomous and context-dependent manner. Indeed, metastability has been conceptually invoked—primarily from a top-down perspective—to interpret diverse biological phenomena: transient intermediate states during sensory organ precursor fate commitment in *Drosophila* [1]; context-dependent identity switching during human cell reprogramming [2]; and primed neural crest states that enable lineage transitions across germ layer boundaries during mammalian development [3]. These examples suggest that biological systems often exploit transiently stable configurations rather than fixed attractors.

Looking forward, an important challenge—and opportunity—lies in constructing cell-autonomous dynamical transition programs using bottom-up synthetic approaches. Such systems would allow metastability to be not merely inferred, but deliberately engineered, thereby offering a powerful lens through which to study how temporal dynamics shape cell fate decisions and enable the ordered assembly of multicellular organisms.

Finally, perhaps the most charming advantage of viewing living matters as dynamical systems is the unifying perspective it affords. Processes that appear superficially unrelated may, at an abstract level, be governed by similar underlying principles. For example, V(D)J recombination during immune cell maturation and receptor expansion along evolutionary lineages both deploy limited genetic building blocks to generate vast diversity through rearrangement and selection, operating at different

timescales. Framed appropriately, such processes may be amenable to description by the same dimensionless equations, differing only in parameter regimes. In this sense, the dynamical-systems viewpoint does not merely explain individual phenomena—it offers a shared theoretical language for understanding biological complexity across scales.

4.2 Future design space: the unpossessed land

As summarized in Chapter 1, synthetic biology has generated an expanding repertoire of tools that have profoundly advanced both biological applications and fundamental understanding. Yet, despite the impressive breadth of existing engineered systems, the design space that remains unexplored is far larger than what has already been charted. With the rapid advances of computational modeling, high-throughput measurements, and AI-assisted design, it is likely that we will soon gain deeper insights into the governing principles underlying biological organization. Against this backdrop, I outline several particularly promising—and largely unpossessed—regions of future design space.

4.2.1 Molecular-level design beyond transcription

At the molecular level, current synthetic toolkits have focused predominantly on transcriptional and post-transcriptional regulation. Engineered transcription factors and small RNAs have enabled programmable control over gene expression with increasing predictability. In contrast, other layers of regulation—such as alternative splicing, post-translational modification, and epigenetic regulation—remain comparatively underexplored. A major bottleneck is that these processes are themselves incompletely understood, limiting their reliable deployment in synthetic contexts.

Alternative splicing illustrates both the scale of opportunity and the depth of challenge. More than 95% of multi-exon genes in the human genome undergo alternative splicing, creating vast potential for rewiring protein interactomes and generating context-dependent functionality [4]. While global factors such as transcriptional elongation rates and splicing regulator abundance influence splicing outcomes [5, 6], alternative splicing is ultimately governed by a diverse set of mechanisms [7]. Existing programmable splicing tools largely co-opt endogenous splicing machinery and rely on naturally derived splice sites [8], which constrains their portability and generalizability across contexts.

A compelling future direction would be the construction of fully orthogonal splicing systems. In such a framework, a single genomic locus could be selectively parti-

tioned into distinct functional domains through programmable splice-site recognition, with splicing decisions directly coupled to defined cellular signals. This would enable cellular events to be recorded or inferred through the resulting mRNA or protein products. Moreover, by mapping specific splicing programs to controllable extracellular inputs, one could systematically probe protein function by dynamically generating defined isoforms—achieving functional diversification through signal modulation rather than repeated genomic integration. Such approaches would shift splicing from a passive source of complexity to an active design dimension.

4.2.2 Functional validation in real biological contexts

At the functional level, many synthetic regulatory elements have been validated primarily using fluorescent reporters in cultured cell lines. Even de novo designed proteins are often characterized only in in-vitro assays. To fulfill the premise of "rewiring life," these components must ultimately be shown to function robustly in physiologically relevant settings.

One illustrative example is the widespread use of the incoherent feedforward loop (IFFL) motif to achieve precise gene expression control, particularly in the context of gene therapy. Although the conceptual appeal of this architecture is well established, its translation into therapeutic settings has been slow. For instance, IFFL-based regulation of MECP2 for Rett syndrome gene therapy has been explored in several studies, yet current results remain preliminary [9, 10]. Parallel efforts using similar circuit logic to control Rai1 expression have so far been tested only in primary mouse cortical neurons [11]. These cases highlight a broader gap between circuit design and in situ functionality. Filling this gap—by demonstrating that engineered motifs can operate reliably within the complexity of living tissues—will be essential for transforming synthetic biology from proof-of-concept demonstrations into deployable biological engineering.

4.2.3 Evolutionary design as an engineering resource

From an evolutionary perspective, emerging experimental and computational tools now make it possible to investigate how ancestral proteins function within defined environments and how they navigate fitness landscapes. Such studies can illuminate the evolutionary trajectories of proteins and, more fundamentally, probe the plasticity of these landscapes themselves.

Beyond their explanatory value, ancient or cross-species proteins may also serve as practical engineering substrates. Their evolutionary distance from modern cellular

networks can render them partially orthogonal, reducing unintended crosstalk with endogenous pathways. For example, a population-control circuit developed in our laboratory exploited the plant hormone auxin as a signaling molecule, precisely because of its minimal interference with mammalian cellular processes [12]. This strategy suggests a broader design principle: evolutionarily distant components can be repurposed as clean signaling channels, expanding the usable design space beyond what is accessible using native molecular parts alone.

Taken together, these molecular, functional, and evolutionary frontiers define a vast unpossessed land for future synthetic biology—one that invites not only new tools, but new ways of thinking about how biological systems can be constructed, interrogated, and repurposed.

4.3 Lessons for 'narrative engineering'

The term *narrative engineering* has been used in the organoid field to describe the ambition of engineering organs not merely as static structures, but as temporally evolving systems shaped by collective, emergent behaviors of multicellular populations. Achieving this goal requires the coordinated integration of multiple layers: cell-autonomous genetic programs or artificially imposed spatial organization; controlled biochemical and mechanical environments that ensure appropriate signaling, nutrient supply, and physical constraints; and external stimuli that guide developmental trajectories over time [13]. In this sense, narrative engineering is fundamentally concerned with *process*, rather than simply outcome.

This perspective closely resonates with the discussion in Chapter 3. Implementing a functional Turing-like pattern *in vivo* is not simply a matter of encoding a reaction–diffusion circuit at the genetic level. It also demands an appropriate 3D tissue context, which may be enabled by advances in 3D printing and biofabrication technologies. Sustained exchange of nutrients and waste is equally essential, often necessitating coordinated metabolic regulation and vascularization around the forming pattern. Yet, the controlled generation of vascular and neural networks surrounding the organoid remains one of the central bottlenecks in the field [14]. Additional challenges arise from growth itself: tissues are dynamic and expanding entities, raising the question of how patterned structures can be preserved, scaled, or remodeled without losing their defining features.

Despite these challenges, synthetic multicellular circuits offer an assumptive narrative framework for the organogenesis. Rather than attempting to directly reproduce

the full complexity of natural development, synthetic circuits articulate a minimal, internally consistent story of how specific regulatory motifs give rise to quantitative behaviors in space and time. In doing so, they define a *prototype*—a generative logic that specifies not only what structures might emerge, but *how* they emerge.

In this regard, synthetic multicellular circuits evoke the idea articulated in C.S.Lewis's The Pilgrim's Regress, in his discussion of "the prototype and the copy." The prototype is not a simplified imitation of reality, but a conceptual original—a form that gives coherence and meaning to its many imperfect realizations. Similarly, a synthetic developmental circuit need not recapitulate every molecular detail of natural development to be valuable. Its power lies in providing a coherent narrative that developmental biologists can reason about, test against reality, and refine.

Looking forward, synthetic circuitry may reshape the field not by replacing classical developmental approaches, but by complementing them. By offering explicit, mechanistically grounded narratives of multicellular organization, synthetic systems can serve as reference frameworks against which natural development is interpreted. In this way, narrative engineering becomes not only a technical endeavor, but an epistemic one: a means of understanding life by learning how to tell its stories in a form that can be built, perturbed, and ultimately understood.

References

[1] Ritika Giri et al. "Single-cell Senseless protein analysis reveals metastable states during the transition to a sensory organ fate". In: *iScience* 25.10 (Oct. 2022), p. 105097.

[2] Kee-Pyo Kim et al. "Donor cell memory confers a metastable state of directly converted cells". In: *Cell Stem Cell* 28.7 (July 2021), 1291–1306.e10.

[3] Antoine Zalc et al. "Reactivation of the pluripotency program precedes formation of the cranial neural crest". In: *Science* 371.6529 (Feb. 2021).

[4] Peter Kjer-Hansen and Robert J Weatheritt. "The function of alternative splicing in the proteome: rewiring protein interactomes to put old functions into new contexts". In: *Nat Struct Mol Biol* 30.12 (Dec. 2023), pp. 1844–1856.

[5] Fangyuan Ding and Michael B Elowitz. "Constitutive splicing and economies of scale in gene expression". In: *Nat Struct Mol Biol* 26.6 (June 2019), pp. 424–432.

[6] Fangyuan Ding et al. "Dynamics and functional roles of splicing factor autoregulation". In: *Cell Rep* 39.12 (June 2022), p. 110985.

[7] Luciano E Marasco and Alberto R Kornblihtt. “The physiology of alternative splicing”. In: *Nat Rev Mol Cell Biol* 24.4 (Apr. 2023), pp. 242–254.

[8] Melina Mathur et al. “Programmable mutually exclusive alternative splicing for generating RNA and protein diversity”. In: *Nat Commun* 10.1 (June 2019), p. 2673.

[9] Paul D Ross et al. “Self-regulating gene therapy ameliorates phenotypes and overcomes gene dosage sensitivity in a mouse model of Rett syndrome”. In: *Sci Transl Med* 17.792 (Apr. 2025), eadq3614.

[10] Michael J Flynn et al. “Synthetic dosage-compensating miRNA circuits allow precision gene therapy for Rett syndrome”. In: *bioRxiv* (Mar. 2024).

[11] Noa Katz et al. “Tunable, proteolytic dosage control of CRISPR-Cas systems enables precise gene therapy for dosage sensitive disorders”. In: *bioRxiv* (Sept. 2025).

[12] Yitong Ma et al. “Synthetic mammalian signaling circuits for robust cell population control”. In: *Cell* 185.6 (Mar. 2022), 967–979.e12.

[13] Takanori Takebe and James M Wells. “Organoids by design”. In: *Science* 364.6444 (June 2019), pp. 956–959.

[14] Ricard Solé et al. “Open problems in synthetic multicellularity”. In: *NPJ Syst Biol Appl* 10.1 (Dec. 2024), p. 151.