The development of RNA-based control systems to regulate signaling and dictate cell fate in a model MAPK pathway

Thesis by

Kate Elizabeth Galloway

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2012

(Defended May 18, 2012)

Acknowledgments

This thesis has been seven years in the making. It represents not only my efforts, but also the efforts of many people who have provided valuable academic, professional, and technical assistance and guidance, not to mention personal support. While this list is not exhaustive, I would like to particularly thank the people below for their much appreciated help.

I am so very grateful to my advisor, Prof. Christina Smolke, who has been a great mentor and inspiration to me. Since I arrived in Christina's lab in early 2006, I have been impressed by her diligence in research and commitment to bringing out the best in her students. Christina's tireless help with editing this manuscript has given me a much deeper appreciation of the nuances of research and the power of language. In my experience, it is rare to find someone who is so excellent in her own work and simultaneously so patient with the development of others. When I think about Christina, I recall a conversation we had in late 2008, only six weeks before our lab's scheduled move to Stanford. I walked into Christina's office to inform her that I couldn't join the lab in moving to Stanford as I had just received the surprising news that I was pregnant, and I needed to stay in Pasadena to keep my family together. Christina never expressed any sentiment other than her concern and unwavering commitment to help me figure out a way to finish my Ph.D. at Caltech. She helped me arrange to remain at Caltech and organized grants and other funding so that I could finish my research here. Additionally, she has spent extra hours on the side coordinating my benefits and pay and meeting both via Skype and in person to continue guiding my work. Throughout all of these challenges, Christina has continued to support my development as a scientist and has

given me great flexibility to be around for my family. I am so grateful to have had Christina as my advisor and for her compassion and guidance throughout the past six years.

I would also like to thank my committee: Prof. Michael Elowitz, Prof. Richard Murray, and Prof. Dave Tirrell. I am particularly thankful to Michael for allowing me to use his lab's Quanta flow cytometer, which was critical for characterizing my system. When the Quanta needed to be moved, Michael willingly released it to its new home in the Murray lab. I am so thankful to Richard for coordinating and financing the move of the Quanta as well as providing space in his lab for it. Thanks to Dave for his willingness to help when I needed equipment in his lab and for serving as the chair of my thesis committee.

The Smolke lab has been a wonderful group of friends and colleagues. I am particularly grateful to Dr. Joe Liang for many encouraging discussions via Google Chat, his efforts as the Stanford-to-Caltech supply train, and his excellent work in developing more stringent switches. My work would not be possible without the new methods and switches that he developed. I am also thankful to Dr. Josh Michener for our biocontrol discussions and for his efforts to keep me awake in seminars. I will remember Leo d'Espaux for modeling work-life balance, practicing the beach entry for the Hermosa Beach triathlon, and sharing his stories of growing up in Cuba, including the one about stroking Fidel Castro's beard as a child. To Dr. Yvonne Chen: I am thankful for your friendship and unparalleled model of diligence and efficiency. I am excited to read about the work that will be coming out of your lab starting in 2013.

The Smolke lab move to Stanford in 2009 gave me the privilege of being adopted by a number of labs over the last several years. I am so thankful to Dr. Anand Asthagiri and the Asthagiri lab for providing me a lab home for two years. I am especially thankful to Dr. Steve Chapman for teaching me his Western blot protocol and sharing his stocks with me. I have been more recently adopted by Dr. Frances Arnold's lab and the German contingent of postdocs. Thanks to Dr. Sabine Bastian for her thoughtfulness in helping me maintain a bench through all the remodeling chaos on the second floor and overcrowding upstairs. I will also remember the reliable Ernie's lunch bunch, Dr. Kersten Rabe, Dr. Ryan Lauchli, Dr. Sebastian Schoof, Dr. Martin Engqvist, and Chris Farwell for our varied discussions ranging from new cloning methods to global politics. The last few years have found me increasingly borrowing time, space, and equipment across the campus. Thanks to the Tirrell, Elowitz, and Murray labs for being welcoming and friendly when I came to borrow things.

To the biocontrol reading group, I am grateful not only for our meetings which broadened my appreciation of synthetic biology, systems biology and modeling, but also for the wonderful people who have contributed so much to me both professionally and personally. I am especially grateful to Dr. Elisa Franco for her help in providing a sounding board for ideas about network architecture that changed my approach to Chapter 3. I am also thankful for her generosity in helping me to develop a model of my system. Thanks also to Ophelia Venturelli for the interesting discussions on all things yeast and particularly the galactose pathway. I look forward to hearing the final word on the systems biology of the galactose network response.

There is a lot of work that goes into administering the grad program at Caltech. I want to say thank you to the people who work behind the scenes on these critical tasks. I am particularly grateful to Kathy Bubash, Laura Lutz-King, and Martha Hepworth. In addition to their work supporting students, staff, and faculty, these three ladies made my transition back to work easier following the births of my children. For IT support, I am thankful to Suresh Gupta for the many times he rescued my computer and even saved the computer for the aforementioned flow cytometer.

I am very thankful for the timely emergence of the unofficial Caltech grad-student-mom's club, whose members include Dr. Katie Brenner, Dr. Melissa Pope, and Dr. Kristin Gleitzman. I am especially glad that I witnessed Katie's pioneering efforts as a grad-student–mom. Her example and encouragement gave me the inspiration and the hope that I needed to complete this journey and without which I am not sure that I would have had the courage to attempt it. Also, I am grateful for Melissa and her friendship as we journeyed together figuring out which doctor to see, how to manage infant feedings, and how to finally get some work done in lab. Sometimes it felt like a wild ride. I am glad I had someone else who was there to share in the unique joys and trials of being a mom in grad school.

I could not have come so far without the help of my family. Each one has helped in their own special way by bringing food, babysitting, or simply emailing me words of encouragement. My parents, Steve and Jane Gropp, have always supported my education, whether it was homeschooling me or driving me 30 miles to high school. They have also been an inspiration to me to work diligently and creatively. I am so grateful for their love and encouragement. My dad tirelessly helped me build my science poster every year and

taught me how to use an X-Acto knife and rubber cement. He instilled his love of design within me. My mom dared to dream big dreams for me and believed I could achieve them long before I could. I hope I have made them proud. I am also grateful for the support of my siblings Anna Johansen and Michael Gropp. Thanks for letting me wear the lab coat and goggles when we were kids and for lots of fun with our "secret chemistry labs", which I am sure horrified Mom. Also, I am so glad for my mother-in-law, Rosemary Galloway, who has helped me finish my thesis by coming and taking care of our family, especially when the children were very little. Additionally, I am thankful to SaraJean Wright, who began as Allie's nanny over two years ago and has since become like family. Thank you for taking excellent care of the babies so I could go to lab and get work done.

Nine years ago, I realized I had met the man I was going to marry. Gordie and I have been on this grad school adventure for all but six weeks of our marriage. What will it be like without grad school? I am sure we will find something to do. I am so grateful for his unwavering support. He has believed in me even when I doubted myself. He has sacrificed his time (and probably some sanity) editing my fellowship applications and research write-ups and patiently listened to whole conversations about thermodynamics, network architectures, and other technical boredom to be by my side at Caltech events. He has coached me through some of the most difficult, scary, and joyous moments that I can remember. He is a wonderful father to our two delightful children, Allie (almost 3 years) and Aaron (almost 3 months). When I entered Caltech, I did not expect to be a mother while in grad school, but I would not wish life to be any other way. Allie and, more recently, Aaron have given me such joy and an ability to enjoy the simple things

through their eyes. Allie recently told me, “Mommy, if you have a dream in your heart, you should get it out.” Over these last seven years, I have realized many of the dreams in my heart, including their births, and now a thesis.

Finally, and most importantly, I am thankful to the one who puts dreams in our hearts to inspire us to strive for these goals: Jesus Christ, creator of the universe and my personal savior. I have been so delighted to look into nature and see the mind of God displayed in the intricately imbedded designs. Moreover, it is miraculous to me that we search out and understand the inner workings of nature and apply that understanding to improve other’s well-being and life here. I am thankful for this beautiful world and for my brief time to work and delight in it, knowing that it is but a shadow of things to come.

Abstract

Cells integrate extracellular information via native signaling pathways to spatially and temporally coordinate complex tasks such as development and the immune response. Cellular programming holds the potential of harnessing the sophisticated and complex biological processes of living cells for diverse applications. In the last decade, cellular reprogramming has emerged as a viable therapeutic strategy. In large, reprogramming strategies have relied on statically programmed levels of gene expression to alter cellular behaviors. To construct more sophisticated programs requires dynamic control of expression and strategies for the facile construction of complex control architectures. Additionally, the application of synthetic programs to the control of native regulatory pathways requires the development of tools for interfacing with these pathways, as well as the construction of stringent controllers. Further, control systems composed of modular and tunable elements will facilitate the expansion of synthetic circuitry to a wide array of natural networks with varying system properties.

Here we describe the development of RNA-based control systems to regulate signaling and dictate cell fate in a model mitogen-activated protein kinase (MAPK) pathway. We construct networks of RNA-based control systems that interface with the *Saccharomyces cerevisiae* mating pathway to dictate entry into one of three programmed alternative fates dependent on environmental stimuli. We present a readily translatable method for identifying control points within natural networks that enable the construction of a modular interface between synthetic circuitry and native networks. In building these networks, we demonstrate the rational tuning of circuit performance via the exchange of well-defined parts to compose networks capable of actuating changes in cellular behavior

in response to environmental cues. Further, we construct network architectures which facilitate reduced interference from simultaneously integrated opposing programs and identified sensitive parameters for engineering robust circuit performance. Finally, we present the development of a novel RNA-based control element for the regulation of both synthetic and endogenous transcripts. This work provides a model for engineering systems that regulate signaling and direct cell fate which may be applied to additional decision-making pathways to advance tissue engineering strategies, treat diseases, and study the behavior of natural regulatory networks.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Cellular differentiation, organism development, and tissue homeostasis require precise temporal and spatial regulation of cellular responses to external inputs. Improper regulation and coordination of these processes can lead to deformities, disease, and death. Controlling cell fate offers the potential to coordinate and redirect cellular trajectories. Efforts to control cell fate have primarily focused on regulating the chemical, biochemical, and mechanical environments in the extracellular space. Yet for a wide range of applications, environmental cues alone are insufficient to alter cellular behaviors. Further, many *in vivo* applications preclude precise control of the extracellular environment. In these instances, effective control strategies must often be performed in antagonistic environments. Synthetic gene regulatory networks that control the internal decision-making process offer an alternative approach to directing cellular fate.

Synthetic gene regulatory networks to control biological systems

The examples of synthetic gene regulatory networks encoding sophisticated functions have steadily increased over the last decade, aided by improved fabrication and sequencing methods that have reduced the cost of cloning [1-5]. Circuits capable of exhibiting dynamic behaviors, processing logical functions, and communicating cellular information have been demonstrated in synthetic gene regulatory networks [6-11]. Complex, multicellular behaviors such as synchronized oscillation and edge detection have been achieved by rational coupling of these various networks [12, 13]. Additionally, building genetic networks from the bottom-up has provided well-defined systems that can be used to study fundamental biological mechanisms [14]. Comparing the performance of

modeled and experimentally realized networks has informed our fundamental understanding of the importance of various process in biological systems, including degradation, cooperativity, and noise [15-17]. Despite remarkable advances in the realization of synthetic circuits, translation of these systems to real-world applications has been limited by the availability of methods to connect them to the requisite information in living cells [18]. There exists a need for modular interfaces that can extract specific information from biological systems and route this information to synthetic gene networks capable of programming rational responses via their interaction with native regulatory networks (Figure 1.1).

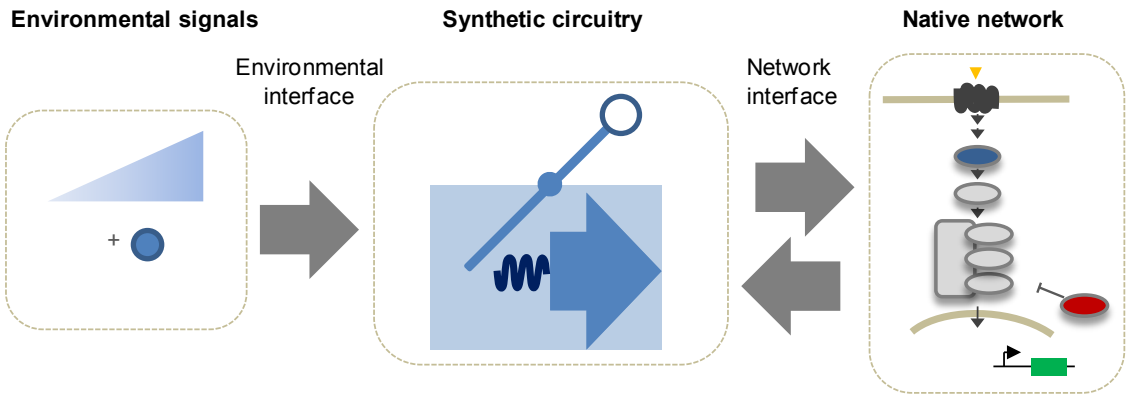


Figure 1.1. Building synthetic circuitry that interfaces with the environment and native regulatory networks to control cellular behavior. Environmental signals are transduced into changes in the native regulatory network via synthetic circuitry. The composition of the synthetic circuitry dictates the environmental interface, how environmental information is input into the synthetic circuit, and the network interface, how the circuitry implements regulation and extracts information from the native network.

Building a modular interface requires that circuitry be composed of modular parts such sensors, actuators, and other components that may be swapped in and out to connect to different environmental cues and various native networks. A synthetic circuit's ability to actuate changes in a native network depends on the properties of the circuit's parts, as well as, their interaction with each other (Figure 1.2A, B). An optimally placed circuit

runs quiescently until activated by the proper environmental cue to trigger a change in the native network (Figure 1.2C). The development of modular parts with tunable properties will enhance design flexibility, facilitating the optimal positioning of synthetic circuits for a broad range of applications.

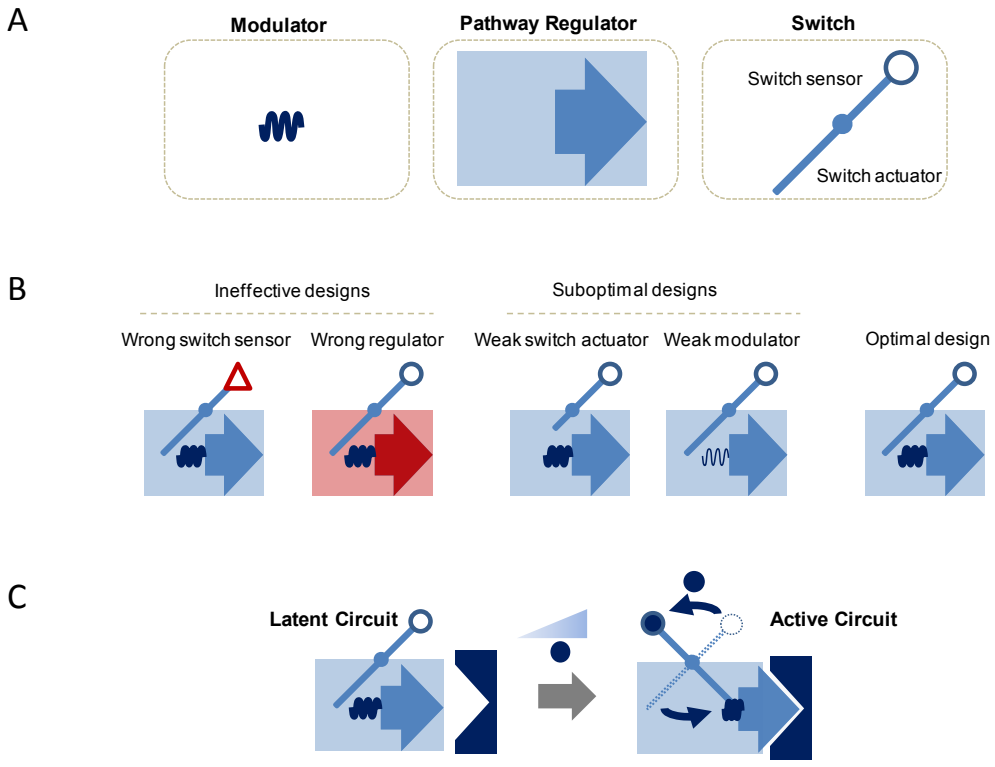


Figure 1.2. Components for interfacing with the environment and native regulatory networks. A. Parts used to interface with the environment and network. Function of the circuit relies on the properties of the elements and how they are integrated into the larger device. **B.** Composing circuits for optimal performance requires selection of the proper parts to efficiently transduce the environmental input into changes in the regulatory machinery. **C.** The latent circuit is activated by increasing environmental signal transduced by a sensor to imbedded synthetic circuitry. The synthetic circuit processes increasing levels of input by raising the profile of regulatory machinery that interfaces with the native network, mediating changes in the native network behavior.

Building a modular interface from ncRNAs

The modular, tunable, and programmable nature of RNA makes it an ideal candidate to perform sensing, actuation, and regulatory functions within a synthetic circuit [19]. Over the last twenty years the scientific community has become increasingly

aware of the role of non-coding RNAs (ncRNAs) in cellular control over gene expression. ncRNAs were thought to provide rather generic functions as ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs) in translation and as small nuclear RNAs (snRNAs) in splicing. Studies of prokaryote genomes and their regulation forged the central dogma of biology, dictating that RNA had a relatively passive role in the transfer of genetic information from genes to proteins. In addition to specifying cellular state, proteins were thought to direct the trajectory of cellular fate by controlling gene expression networks. The vast tracks of ncRNA found in eukaryotes were hypothesized to be the evolutionary accumulation of inert sequences. However, large sets of these sequences are transcribed [20]. In fact, the majority of transcribed sequences in higher eukaryotes are never translated [21]. Additionally, the percentage of non-coding transcripts scales with organism complexity, while the number of coding genes does not [22]. With increasing discoveries of small ncRNAs that modulate gene expression [23-25] as well as with the discovery of the RNA interference (RNAi) pathway [26-28], ncRNAs appear to be the defining layer of sophisticated biological control that differentiates species with remarkably similar genomes [29].

While prokaryotes are known to use ncRNA to regulate gene expression, their dominant layer of control relies on protein-based regulation of transcription. Therefore, it is not surprising that prokaryotes have a relatively diverse genome and proteome from which to select elements to serve as regulators of gene expression. In contrast, eukaryotic organisms rely on a diversity of ncRNA to compensate for a largely stable and relatively small genome. As suggested by eukaryotic gene regulation strategies, control schemes for modulating gene expression are likely to require additional layers of control including the

implementation of RNA based-control systems that provide post-transcriptional control. To date, synthetic network engineering has largely focused on protein-based regulatory elements to build control systems [30]. However, there is a fundamental limitation to the complexity and specificity of protein-based regulatory schemes, which have a higher energetic cost. The burden of carrying additional genes encoding single protein regulators increases exponentially as complexity is introduced. Ultimately, in these accelerating networks, the cost of producing another protein regulator exceeds the benefit to the organism [31]. Regulation through alternative mechanisms such as ncRNA offers organisms an alternative that allows for specificity and diversity at a lower energetic cost [32].

The recurrence of ncRNA as a dominate regulatory motif in higher eukaryotes suggests that ncRNA-based regulators may improve the construction of complex regulatory architectures in synthetic circuits and facilitate connections between synthetic and endogenous circuitry. Additionally, the tunable properties of RNA and the potential to predict structure stability through existing algorithms make RNA a powerful substrate on which to build synthetic gene expression control systems. While post-transcriptional control of gene expression generally exhibits modest regulatory activities compared to transcriptional control, placement of post-transcriptional regulators at sensitive control points can mediate significant effects [33, 34]. Further, moderate activities endow these regulators with additive effects that provide a robust mechanism for tuning gene expression [35] and buffering the noise of gene expression in natural systems [36]. Finally, the additive nature of ncRNAs with modest activities provides a mechanism for synthetic circuits to logically process multiple inputs to orchestrate rational responses

[37]. In composing synthetic circuits, ncRNAs provide an additional degree of freedom for tuning circuit performance within the desired application.

Selecting for RNA-based sensors

To construct a modular environmental interface, requires modular sensors that can transduce the environmental input into changes in the synthetic circuit. We propose that RNA aptamers demonstrate the requisite modularity as environmental sensors that can be wired to actuation and regulatory elements. Aptamers are a class of small nucleic acids, including some ncRNAs, that bind to a wide range of ligands, such as small molecules and peptides, with sensitivity and selectivity that can rival that of proteins [38]. Aptamers are thought to bind ligands through a process called adaptive recognition, in which ligand binding occurs as the RNA molecule transitions through relatively unstructured conformations until the appropriate binding pocket is formed. Upon formation of the binding pocket, the aptamer associates with the ligand which stabilizes the ligand-bound structure. Due to evolutionary pressure during selection, the three-dimensional structure of aptamer complexes reflects highly optimized scaffolds for ligand recognition [39].

The development of new aptamer sequences to cellular molecules of interest offers the potential to connect to endogenous networks in a rational way that can direct information into exogenous control systems. Synthetic RNA aptamers have been generated *de novo* to various small-molecule and protein targets through *in vitro* selection or SELEX strategies [40, 41]. Briefly, a large library of RNA molecules ($\sim 10^{14}$–10^{15}) is incubated with the target of interest. Functional aptamers within this library space are subsequently partitioned from nonfunctional members and collected, typically using an affinity chromatography based separation strategy.

Collected sequences are then reverse transcribed and amplified to generate an enriched library that will serve as the input pool for the next round of selection. The Smolke laboratory has generated RNA aptamers that exhibit varying specificities to benzylisoquinoline alkaloids [42] and folinic acid derivates, and is developing high-throughput strategies for the direct selection and characterization of new protein- and small-molecule-responsive aptamers. Developing modular interfaces that facilitate information exchange between natural and engineered systems is critical for constructing biological control systems that program cell behavior.

Natural RNA switches as gene expression control systems

Naturally occurring RNA switches called riboswitches have been shown to regulate gene expression in response to a variety of metabolites, constructing various metabolic feedback control loops [43, 44]. Riboswitches are naturally occurring *cis*-acting RNA regulatory elements that modulate gene expression events in response to changes in intracellular metabolite concentrations [43, 44]. Riboswitches are comprised of at least two functional domains: a sensor or metabolite-binding domain and an actuator domain. Metabolite binding to the riboswitch occurs as the RNA molecule surveys equilibrium conformations. Once the appropriate binding pocket is formed, the ligand can bind the sensor domain. Ligand binding biases the equilibrium of the sensor region toward ligand-bound conformations. Through linker modules, conformational changes in the sensor domain induce a conformational change in the actuator element. In a riboswitch, the actuator is called an expression platform because it regulates gene expression [45]. Gene regulation occurs as metabolite binding events shift the equilibrium of the riboswitch toward its active or inactive conformation in which the expression platform adopts an active or inactive functional state. When the expression platform

adopts its active conformation, gene regulation occurs through an array of diverse mechanisms such as transcription termination, mRNA cleavage, or translation initiation [19].

Riboswitches are implemented by the cell as autonomous biological control systems. These RNA elements provide feedback control by sensing metabolites that are substrates and products of the riboswitch-regulated enzymes and modulating the levels of these enzymes in response to cellular metabolite concentrations. One such example is the glutamine-fructose-6-phosphate (GlcN6P) amidotransferase ribozyme-based riboswitch that is located within the 5' untranslated region (UTR) of the *glmS* gene. This enzyme metabolizes GlcN6P, which is the small-molecule effector of the riboswitch located upstream of *glmS* [46]. While some riboswitches have been discovered to promote gene expression [47], the majority repress the expression of their target gene. However, these natural RNA switches generally have evolved nonmodular architectures, in which the sequences of the sensor and actuator components interact to allow the switch to adopt different functional conformations, making the adaptation of these ncRNA controllers to new molecular inputs and regulatory mechanisms through direct component swapping unfeasible [19, 45, 47].

Synthetic RNA switches that act through ribozyme-based cleavage mechanisms

Recent studies have demonstrated the design and implementation of synthetic riboswitch counterparts by pairing RNA aptamers with various ncRNA expression platforms [34, 48, 49]. The ncRNA regulatory element encoded in the actuation domain dictates the output flexibility of the controller, in terms of the genetic targets that can be regulated, and the organisms in which the controller functions. Therefore, regulatory elements that exhibit function in diverse organisms and that can be used to flexibly

regulate diverse genetic targets are of interest for integration into these synthetic controllers. A number of design strategies have been developed to functionally couple aptamers to small molecules and proteins to diverse ncRNA regulatory elements, from miRNAs [48] to ribozymes [50], such that binding of the ligand to the aptamer domain results in a change in the activity of the ncRNA regulatory element.

Ribozymes are RNA molecules that catalyze a variety of reactions such as self-cleavage or ligation [51]. Thus, ribozyme activity is independent of cell-specific machinery, and these RNA elements may provide a regulatory strategy that can be used across diverse organisms, including bacteria and eukaryotic cells. The hammerhead ribozyme is one of the most extensively studied ribozymes [51-54]. Previous work coupled aptamers to the stem-loop regions in hammerhead ribozymes, allowing for *in vitro* allosteric ribozymes [55-57]. However, the coupling strategies used in these early allosteric ribozyme designs inactivated the ribozyme activity at physiological salt conditions, not allowing these switches to be implemented as controllers inside cells. Through elucidation of the design rules for *in vivo* catalytic activity [58, 59], the stage was set for the design of RNA switch platforms that functionally integrate hammerhead ribozymes as *in vivo* regulatory elements.

The Smolke laboratory recently demonstrated a modular and extensible framework for engineering *in vivo* ligand-regulated ribozyme switches. The described switch device contains three distinct functional domains: sensor domain, comprised of an aptamer sequence, a transmitter domain, and actuator domain, comprised of a hammerhead ribozyme sequence (Figure 1.3) [60]. When placed in the 3' untranslated region (3' UTR), the engineered ribozymes switches regulate gene expression. In the OFF

state, the ribozyme favors the active conformation, promoting cleavage of the transcript. Addition of ligand biases the ribozyme switch to the ligand-bound inactive conformation, resulting in increased gene expression. Small-molecule-dependent regulation of gene expression has been demonstrated on various heterologous genes and enabled the construction of RNA switches exhibiting up- and down-regulation of target expression levels. The design of the transmitter domain is a critical design feature that supports the insulation and modularity of the sensor and actuator components and allows the forward design and tuning of synthetic RNA switches. The resulting ability to mix-and-match sensing and actuation domains makes the modular RNA switch platforms powerful tools for developing tailored gene expression control systems.

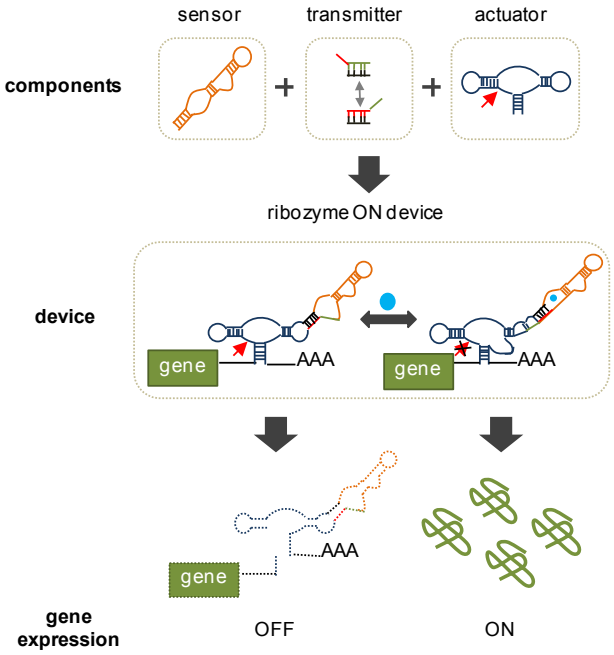


Figure 1.3. RNA-based switches regulate gene expression in response to small-molecule concentration. The switch device is constructed from three primary components, a sensor, a transmitter, and an actuator. These devices regulate the expression of a gene when placed in the 3' untranslated region (3' UTR). Shown above in the OFF state, the switch reduces gene expression in the absence of the small molecule by cleaving and destabilizing the transcript. Presence of the ligand alters the structure of the sensor which is transduced to the changes in the actuator structure via the transmitter. In the presence of ligand, the switch is ON and the actuator structure allows for increased gene expression. Adapted from Liang, J, et al. (*Submittedn*).

The construction of ligand-responsive control elements that act in *trans* will improve our ability to study natural gene networks as well as impose exogenous control in biological systems. The described ribozyme switch system allows for the control of transgenes (in cis); however, it cannot be directly used for the regulation of endogenous genes (in trans). While a number of synthetic RNA switches that allow for the regulation of endogenous genetic targets have been described that function through miRNA-[48], shRNA- [61], or antisense-based [62] mechanisms, these regulatory elements are associated with a number of drawbacks. For example, antisense-based regulatory efficiency is highly variable across targets [63]. In addition, utilizing the RNAi pathway to process synthetic substrates has raised concerns about off-target effects and competition between synthetic and native substrates [64]. Cis-acting ribozymes may be converted into to trans-acting ribozymes by splitting the stem I loop and engineering the intermolecular reaction between the ribozyme and target sequence (Figure 1.4). Facilitating the ribozyme-target binding event *in vivo* is a critical hurdle to implementing trans-acting ribozymes as regulators of gene expression [65]. Previous attempts to control target expression via trans-ribozyme have been limited by poor *in vivo* efficiency [66, 67]. Improvements in ribozyme designs and expression systems may increase ribozyme-mediated knockdown of targets *in vivo.* Elucidating the rules for the regulation of target transcripts will poise trans-acting ribozymes as unique actuators of gene expression. Finally, extension of the design principles elucidated for functional *in vivo* activity of cis-ribozyme switches to the design of trans-ribozyme switches may provide a promising alternative to target the expression of endogenous proteins without modifying the natural context of the target gene [66, 68, 69].

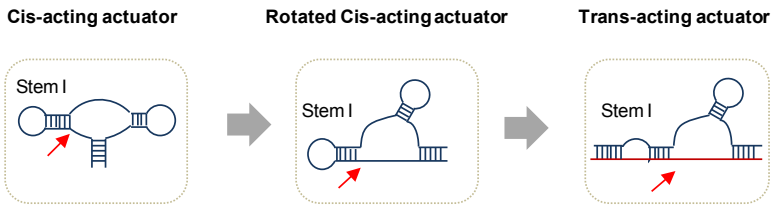


Figure 1.4. Converting cis-acting actuators to trans-acting requires engineering an intramolecular reaction into an intermolecular reaction. The cis-acting ribozyme (at left) can be converted to a trans-acting ribozyme by opening up stem I of the ribozyme to allow binding of a target transcript (shown in red).

MAPK cascades as universal signaling modules in eukaryotes

Mitogen-activated protein kinase (MAPK) cascades are highly conserved signaling pathways that control such processes as differentiation, mitosis, and apoptosis (Figure 1.5) [70]. Signaling through this pathway begins when extracellular signals are transduced across the cell membrane through receptor binding events that activate G-proteins. G-proteins relay these signals by facilitating phosphorylation of MAPKKKKs that continue phosphorylation through a three-tiered cascade until reaching the MAPK [71]. MAPKs regulate cellular behavior through interaction with repressors and transcription factors that determine entry into various cellular programs [72]. Many human cancers and other diseases are known to result from aberrant activation of cellular programs connected to MAPK signaling [73, 74]. Thus, controlling improperly activated signals has important implications in the development of therapeutics. Control systems that modulate MAPK pathways will provide an opportunity to interface with a large class of endogenous regulatory networks by which cellular fate can be programmed.

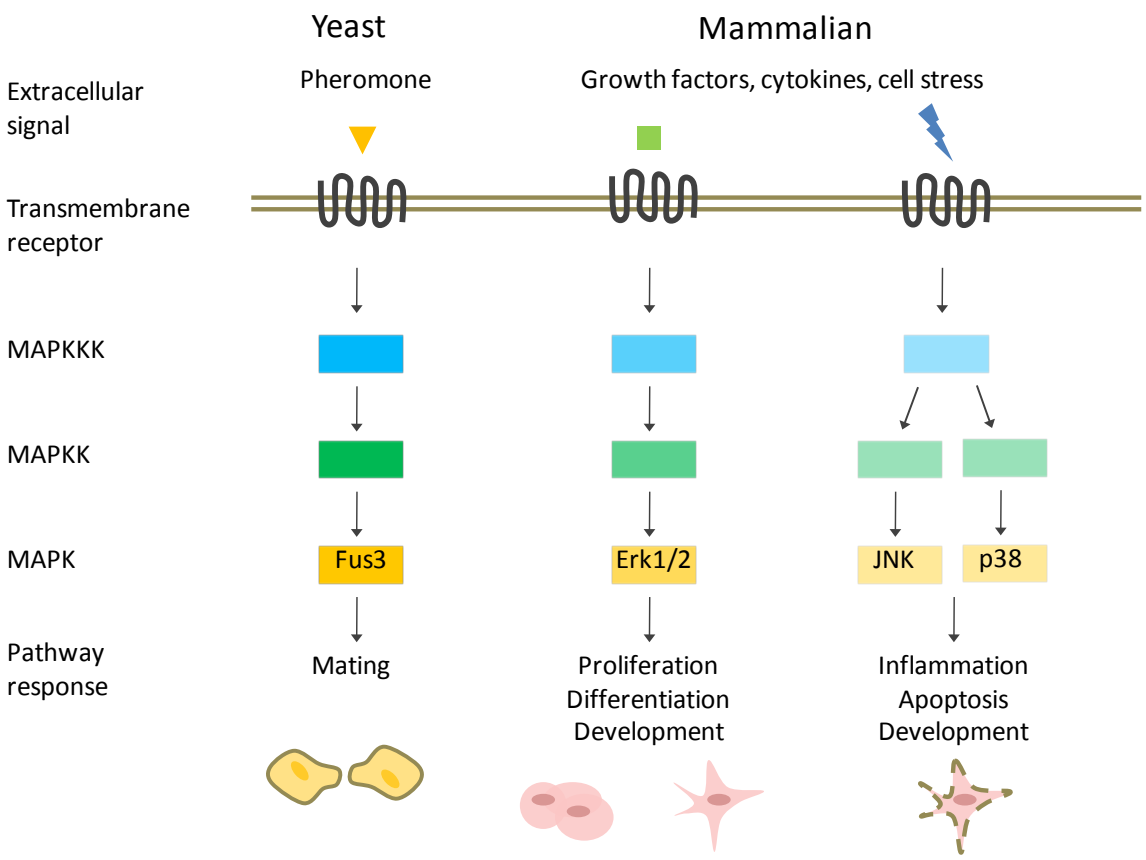


Figure 1.5. MAPK cascades as universal signaling modules in eukaryotes. From yeast to mammals MAPK pathways preserve the transmembrane receptor and downstream three-tiered MAPK cascade that ultimately generates a phenotypic response to the stimuli. Adapted from [75].

Despite an increase in our understanding of MAPK cascades, the development of therapeutics to intervene and redirect cellular fate through targeting components of this pathway has been primarily limited to kinase inhibitors [76-78]. These inhibitors act competitively to limit signal transduction; however, in pathways where control loops provide redundant verification of signaling, these inhibitors may be overwhelmed. Additionally, the delivery of these inhibitors is not restricted to diseased cells, which can result in unintended toxic side effects in healthy cells [79]. Effective therapeutics that redirect aberrant signaling through these pathways may need to compete with transcriptional feedback and discriminate between healthy and diseased cells. Implementation of control systems that regulate protein

levels offers the potential to rationally mediate MAPK signaling. Control systems that dictate MAPK signaling will provide a tool to elucidate our understanding of these pathways and potentially serve as a therapeutic strategy to counteract aberrant activation of cellular programs.

The homology between MAPK cascades in single-celled organisms such as yeast and higher eukaryotes allows for comparison of parallel pathway responses [71]. In particular, the study of the *Saccharomyces cerevisiae* MAPK pathways has illuminated paradigms in human MAPK signaling such as signal insulation through scaffold proteins [80, 81]. In the model eukaryotic organism *S. cerevisiae*, multiple MAPK cascades direct cellular fate via divergent regulatory programs. Decisions to halt cell cycle, upregulate excretion of a metabolite, or change cell morphology are programmed as the rational response to environmental signals. *S. cerevisiae* responds to pheromone by activating a receptor-coupled-G-protein three-tiered MAPK cascade (Figure 1.6A) [82]. The signal is transmitted from the G-proteins to the MAPK cascade of Ste11, Ste7, and Fus3. Fus3 translocates to the nucleus and phosphorylates the transcription factor Ste12. Fus3 is deactivated by the phosphatase Msg5. Upon pheromone stimulation, cells undergo cell cycle arrest, as can be seen in a halo assay, perform polarized growth to adopt the shmoo morphology (Figure 1.6B), and increase expression from the Fus1 promoter (pFUS1).

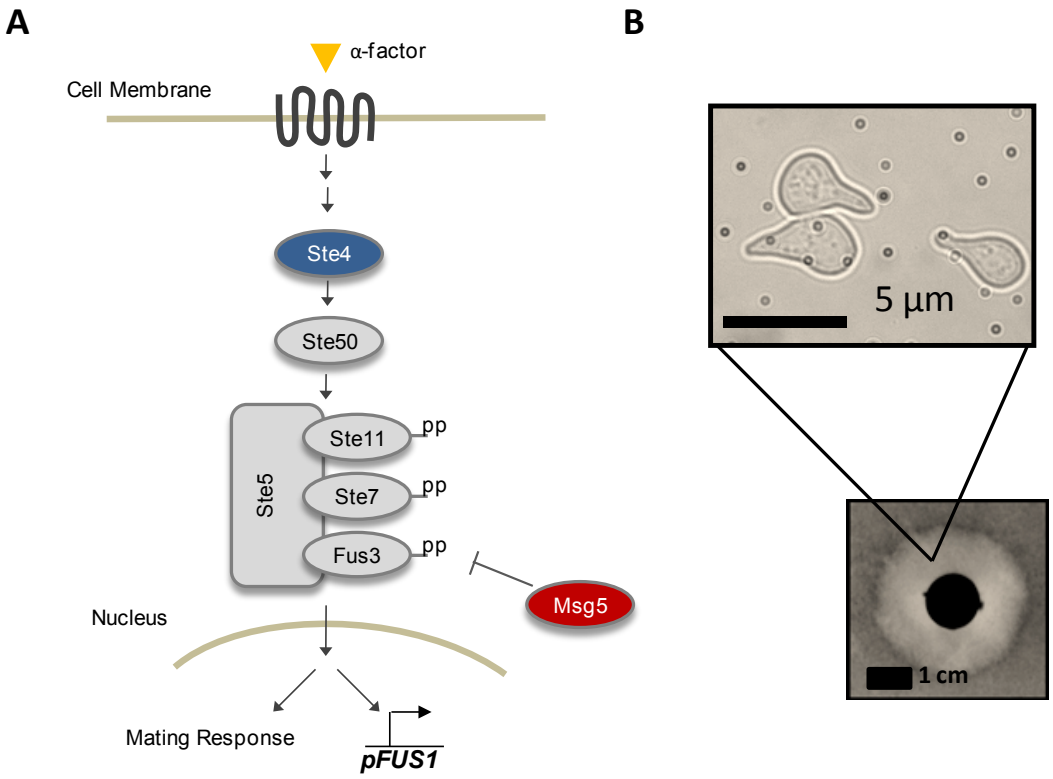


Figure 1.6. Yeast mating pathway and phenotypic response. A. Pheromone (α-factor) binding the transmembrane receptor initiates signaling in the internal G-proteins which is relayed to the canonical three-tiered MAPK cascade (Ste11, Ste7, Fus3). Phosphorylation is relayed down the cascade and culminates with phosphorylated Fus3 translocating to the nucleus to activate a range of transcription factors, transcription at mating genes, and ultimately the canonical mating response. Signaling is antagonized by Msg5, a phosphatase specific to Fus3. **B.** Phenotypic evaluation of the mating pathway can be performed via halo assay and by observing cell morphology. At bottom, a typical halo assay with a filter paper (center dark circle) saturated with pheromone establishing a gradient of pheromone. Cells within a particular radius corresponding to a particular concentration undergo pheromone-induced cell cycle arrest generating a halo in which cell density is significantly reduced relative to the plate outside of this radius. Above, cells stimulated with pheromone form "shmoos" by undergoing polarized cell growth.

Combining synthetic biology and systems biology to build gene regulatory networks that control cellular fate

Several modeling efforts have focused on evaluating the dynamics of MAPK cascades and posited that levels of signaling molecules are responsible for divergent cell fates [83-85]. Experimental results in the yeast pheromone-responsive MAPK pathway have demonstrated that particular profiles of signaling molecules are associated with

entry into specific fates [86-88]. Further, it has been suggested that network topology and the associated positive and negative feedback loops are ultimately responsible for these profiles and thus phenotype [87, 89]. Recent work in mammalian PC-12 cells indicates that the induced ERK MAPK network topology and resulting dynamics direct cell fate (Figure 1.7) [90]. Altering network topology routes cells to alternative fates. Questions still remain as to whether varying the induced network topology represents a conserved strategy across multiple MAPK cascades and eukaryotic organisms. Nevertheless, these results bode well for employing synthetic layers of positive and negative feedback loops as an engineering strategy to regulate cellular behavior

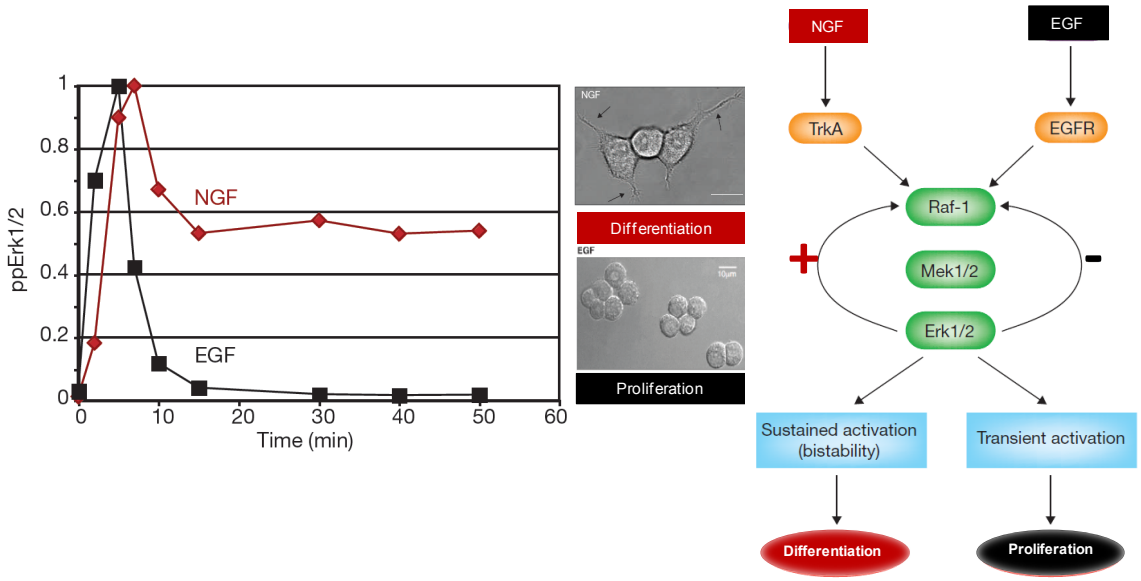


Figure 1.7. Induced network topology shapes the dynamics in a natural regulatory pathway dictating cellular fate. A. Transient stimulation with NGF and EGF leads to different Erk1/2 profiles that correspond to divergent cell fates. Adapted from [90]. **B.** Graph of network topology induced by different growth factors establishes feedback loops of different signs leading to divergent cell fates. Adapted from [91].

Construction of synthetic biological control systems that interact with natural circuits has seen success in regulating pathway activity by incorporating endogenous promoters in feedback control schemes [81]. Additionally, construction of protein

scaffold chimeras has routed cells to an alternative MAPK response [92]. While this work has successfully modulated pathway activity and/or fate, these strategies primarily rely on genetic knockouts of endogenous genes. We prefer a less invasive scheme with a paucity of genetic manipulations to the host that may minimize difficulty in transferring these control strategies to higher eukaryotes. Identification of control points within the molecular network which are sensitive to exogenous control systems will facilitate the construction of noninvasive control strategies. Modulation of the expression of pathway components at these control points can reshape the network response and redirect cellular fate. Further, constructing feedback loops at these control points can fundamentally reshape network topology, alter dynamic signaling profiles, and enhance the robustness of phenotypic selection [90, 93]. Layering these exogenous control systems with RNA-based controllers offers the potential to exogenously induce network topologies that redirect cell fate. We term these synthetic, exogenous control systems "molecular network diverters" as they conditionally divert the molecular network and consequently route cell fate. Molecular network diverters provide a means for orthogonally controlling cell fate within a genetically homogenous population via exogenously applied small-molecule input. Orthogonal control via diverters provides an additional degree of freedom in specifying cell fate preserving existing mechanical, chemical, and biochemical channels for directing cell fate. Molecular network diverters may facilitate the *ex vivo* construction of complex tissues from progenitor cells with imbedded exogenous control systems. These systems may guide progenitor cells to develop normal tissues when seeded on designer scaffolds by supplementing the missing boundary conditions normally present during *in vivo* development and serve to complement traditional tissue engineering

approaches (Figure 1.8). Such systems may be realized in the near future as researchers continue to unravel the systems biology governing cell-fate decisions [94, 95]. Advances in gene therapy delivery may allow molecular network diverters to be translated *in vivo* as cancer therapeutics targeting hyperactive MAPK pathways. Finally, the selection of new sensors responsive to pathway components may allow these diverters to perform autonomous corrective control of cell fate.

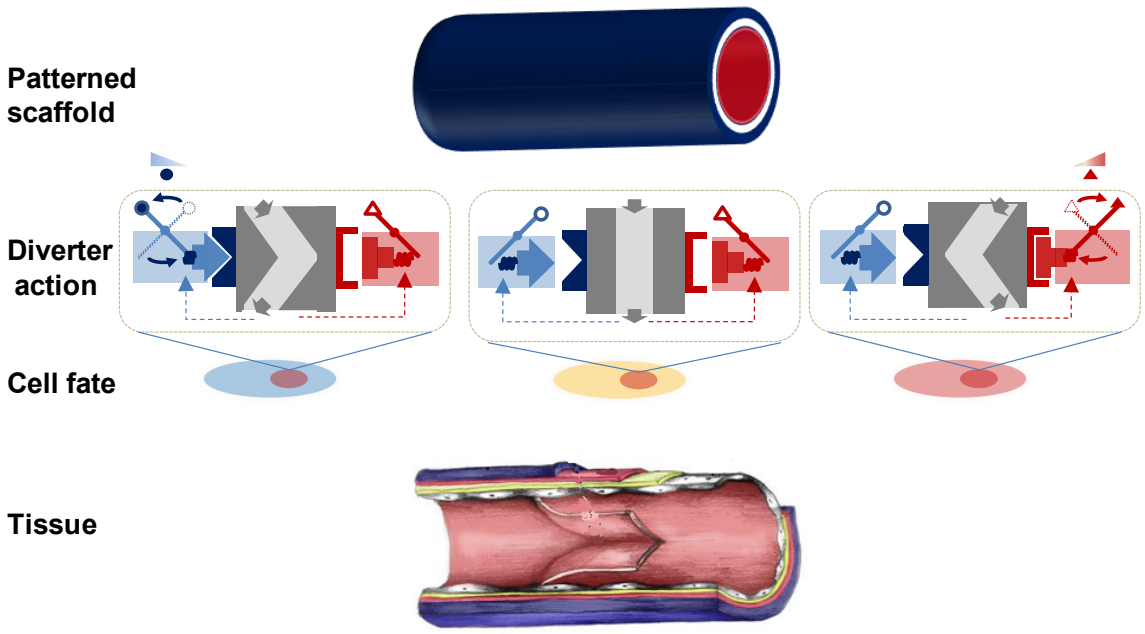


Figure 1.8. Potential application of molecular network diverters to tissue engineering via small-molecule regulated patterning of cell fate. A scaffold with the appropriate geometry is patterned with two small-molecules to trigger diverter action at particular regions within the scaffold. Activated diverters route cells to two alternative fates. All three fates are properly distributed to compose the constructed tissue.

Thesis organization

This thesis is organized into two primary sections. The first section focuses on constructing RNA-based control systems that regulate signaling in the yeast mating pathway. Chapter 2 focuses on using a synthetic titration system to identify regulators of

pathway activity and tracing the pathway response curve that routes cells to alternative fates via varying regulator expression. Using this knowledge, we construct synthetic circuits called “molecular network diverters” composed with engineered RNA controllers and feedback modules that conditionally route cellular fate. Chapter 3 examines the construction of diverters with more complex network architectures composed of multiple modules with different expression modes and RNA controllers that amplify ligand-induced phenotypic switching. We demonstrate an integrated network diverter capable of routing genetically identical cells to one of three fates dependent on environmental signals received. In the second section, chapter 4 discusses RNA-based controllers and efforts to develop a ligand-responsive trans-ribozyme platform that may be used to target both heterologous and synthetic transcript enhancing the design flexibility of synthetic control systems.

As systems biology unravels the inner workings of natural molecular networks, synthetic biology is developing the genetic regulatory tools to implement control systems that guide, tune, and override endogenous network responses. In this work, we utilize principles and tools from both systems biology and synthetic biology to compose a modular and tunable model control system by which we can conditionally direct cell fate. As the array of tools for controlling systems expands, proof-of-principle systems such as ours may be extended to applications in tissue engineering, therapeutics, and beyond.

References

1. Purnick PE, Weiss R: **The second wave of synthetic biology: from modules to systems.** *Nat Rev Mol Cell Biol* 2009, **10:**410-422.
2. Endy D: **Foundations for engineering biology.** *Nature* 2005, **438:**449-453.
3. Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, Benders GA, Montague MG, Ma L, Moodie MM, et al: **Creation of a bacterial cell controlled by a chemically synthesized genome.** *Science* 2010, **329:**52-56.
4. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO: **Enzymatic assembly of DNA molecules up to several hundred kilobases.** *Nat Methods* 2009, **6:**343-345.
5. Canton B, Labno A, Endy D: **Refinement and standardization of synthetic biological parts and devices.** *Nat Biotechnol* 2008, **26:**787-793.
6. Elowitz MB, Leibler S: **A synthetic oscillatory network of transcriptional regulators.** *Nature* 2000, **403:**335-338.
7. Stricker J, Cookson S, Bennett MR, Mather WH, Tsimring LS, Hasty J: **A fast, robust and tunable synthetic gene oscillator.** *Nature* 2008, **456:**516-519.
8. Tigges M, Marquez-Lago TT, Stelling J, Fussenegger M: **A tunable synthetic mammalian oscillator.** *Nature* 2009, **457:**309-312.
9. Rinaudo K, Bleris L, Maddamsetti R, Subramanian S, Weiss R, Benenson Y: **A universal RNAi-based logic evaluator that operates in mammalian cells.** *Nat Biotechnol* 2007, **25:**795-801.

10. Tamsir A, Tabor JJ, Voigt CA: **Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'.** *Nature* 2011, **469:**212-215.

11. Win MN, Smolke CD: **Higher-order cellular information processing with synthetic RNA devices.** *Science* 2008, **322:**456-460.

12. Danino T, Mondragon-Palomino O, Tsimring L, Hasty J: **A synchronized quorum of genetic clocks.** *Nature* 2010, **463:**326-330.

13. Tabor JJ, Salis HM, Simpson ZB, Chevalier AA, Levskaya A, Marcotte EM, Voigt CA, Ellington AD: **A synthetic genetic edge detection program.** *Cell* 2009, **137:**1272-1281.

14. Elowitz M, Lim WA: **Build life to understand it.** *Nature* 2010, **468:**889-890.

15. Eldar A, Elowitz MB: **Functional roles for noise in genetic circuits.** *Nature* 2010, **467:**167-173.

16. Hasty J, McMillen D, Collins JJ: **Engineered gene circuits.** *Nature* 2002, **420:**224-230.

17. Rosenfeld N, Elowitz MB, Alon U: **Negative autoregulation speeds the response times of transcription networks.** *J Mol Biol* 2002, **323:**785-793.

18. Kobayashi H, Kaern M, Araki M, Chung K, Gardner TS, Cantor CR, Collins JJ: **Programmable cells: interfacing natural and engineered gene networks.** *Proc Natl Acad Sci U S A* 2004, **101:**8414-8419.

19. Isaacs FJ, Dwyer DJ, Collins JJ: **RNA synthetic biology.** *Nat Biotechnol* 2006, **24:**545-554.

20. Mattick JS: **Non-coding RNAs: the architects of eukaryotic complexity.** *EMBO Rep* 2001, **2:**986-991.

21. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C, et al: **The transcriptional landscape of the mammalian genome.** *Science* 2005, **309:**1559-1563.

22. Taft RJ, Pheasant M, Mattick JS: **The relationship between non-protein-coding DNA and eukaryotic complexity.** *Bioessays* 2007, **29:**288-299.

23. Moss EG: **Non-coding RNA's: lightning strikes twice.** *Curr Biol* 2000, **10:**R436-439.

24. Meister G, Landthaler M, Dorsett Y, Tuschl T: **Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing.** *Rna* 2004, **10:**544-550.

25. Meister G, Tuschl T: **Mechanisms of gene silencing by double-stranded RNA.** *Nature* 2004, **431:**343-349.

26. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC: **Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.** *Nature* 1998, **391:**806-811.

27. Dykxhoorn DM, Lierberman J: **The Silent Revolution: RNA Interference a Basic Biology, Research Tool, and Therapeutic.** *Annual Review of Medicine* 2005, **56:**401-423.

28. Hannon GJ, Rossi JJ: **Unlocking the potential of the human genome with RNA interference.** *Nature* 2004, **431:**371-378.

29. Mattick JS: **RNA regulation: a new genetics?** *Nat Rev Genet* 2004, **5:**316-323.

30. van Nimwegen E: **Scaling laws in the functional content of genomes.** *Trends Genet* 2003, **19:**479-484.

31. Mattick JS, Gagen MJ: **Mathematics/computation. Accelerating networks.** *Science* 2005, **307:**856-858.

32. Mattick JS, Gagen MJ: **The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms.** *Mol Biol Evol* 2001, **18:**1611-1630.

33. Chen YY, Jensen MC, Smolke CD: **Genetic control of mammalian T-cell proliferation with synthetic RNA regulatory systems.** *Proc Natl Acad Sci U S A* 2010, **107:**8531-8536.

34. Culler SJ, Hoff KG, Smolke CD: **Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins.** *Science* 2010, **330:**1251-1255.

35. Pfleger BF, Pitera DJ, Smolke CD, Keasling JD: **Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes.** *Nat Biotechnol* 2006, **24:**1027-1032.

36. Osella M, Bosia C, Cora D, Caselle M: **The role of incoherent microRNA-mediated feedforward loops in noise buffering.** *PLoS Comput Biol* 2011, **7:**e1001101.

37. Xie Z, Wroblewska L, Prochazka L, Weiss R, Benenson Y: **Multi-input RNAi-based logic circuit for identification of specific cancer cells.** *Science* 2011, **333:**1307-1311.

38. Navani NK, Li Y: **Nucleic acid aptamers and enzymes as sensors.** *Curr Opin Chem Biol* 2006, **10:**272-281.

39. Hermann T, Patel DJ: **Adaptive recognition by nucleic acid aptamers.** *Science* 2000, **287:**820-825.

40. Tuerk C, Gold L: **Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase.** *Science* 1990, **249:**505-510.

41. Bowser MT: **SELEX: just another separation?** *Analyst* 2005, **130:**128-130.

42. Win MN, Klein JS, Smolke CD: **Codeine-binding RNA aptamers and rapid determination of their binding constants using a direct coupling surface plasmon resonance assay.** *Nucleic Acids Res* 2006, **34:**5670-5682.

43. Tucker BJ, Breaker RR: **Riboswitches as versatile gene control elements.** *Curr Opin Struct Biol* 2005, **15:**342-348.

44. Bauer G, Suess B: **Engineered riboswitches as novel tools in molecular biology.** *J Biotechnol* 2006, **124:**4-11.

45. Mandal M, Breaker RR: **Gene regulation by riboswitches.** *Nat Rev Mol Cell Biol* 2004, **5:**451-463.

46. Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR: **Control of gene expression by a natural metabolite-responsive ribozyme.** *Nature* 2004, **428:**281-286.

47. Winkler WC, Breaker RR: **Genetic control by metabolite-binding riboswitches.** *Chembiochem* 2003, **4:**1024-1032.

48. Beisel CL, Chen YY, Culler SJ, Hoff KG, Smolke CD: **Design of small molecule-responsive microRNAs based on structural requirements for Drosha processing.** *Nucleic Acids Res* 2011, **39:**2981-2994.

49. Babiskin AH, Smolke CD: **Engineering ligand-responsive RNA controllers in yeast through the assembly of RNase III tuning modules.** *Nucleic Acids Res* 2011, **39:**5299-5311.

50. Win MN, Smolke CD: **A modular and extensible RNA-based gene-regulatory platform for engineering cellular function.** *Proc Natl Acad Sci U S A* 2007.

51. Long DM, Uhlenbeck OC: **Self-cleaving catalytic RNA.** *Faseb J* 1993, **7:**25-30.

52. Pley HW, Flaherty KM, McKay DB: **Three-dimensional structure of a hammerhead ribozyme.** *Nature* 1994, **372:**68-74.

53. Hammann C, Norman DG, Lilley DM: **Dissection of the ion-induced folding of the hammerhead ribozyme using 19F NMR.** *Proc Natl Acad Sci U S A* 2001, **98:**5503-5508.

54. Blount KF, Uhlenbeck OC: **The structure-function dilemma of the hammerhead ribozyme.** *Annu Rev Biophys Biomol Struct* 2005, **34:**415-440.

55. Soukup GA, Breaker RR: **Nucleic acid molecular switches.** *Trends Biotechnol* 1999, **17:**469-476.

56. Soukup GA, Breaker RR: **Engineering precision RNA molecular switches.** *Proc Natl Acad Sci U S A* 1999, **96:**3584-3589.

57. Soukup GA, Breaker RR: **Allosteric nucleic acid catalysts.** *Curr Opin Struct Biol* 2000, **10:**318-325.

58. De la Pena M, Gago S, Flores R: **Peripheral regions of natural hammerhead ribozymes greatly increase their self-cleavage activity.** *Embo J* 2003, **22:**5561-5570.

59. Khvorova A, Lescoute A, Westhof E, Jayasena SD: **Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity.** *Nat Struct Biol* 2003, **10:**708-712.

60. Win M, Smolke C: **Universal riboswitch platforms for engineering cellular function.** In Submission.

61. Beisel CL, Bayer TS, Hoff KG, Smolke CD: **Model-guided design of ligand-regulated RNAi for programmable control of gene expression.** *Mol Syst Biol* 2008, **4:**224.

62. Bayer TS, Smolke CD: **Programmable ligand-controlled riboregulators of eukaryotic gene expression.** *Nat Biotechnol* 2005, **23:**337-343.

63. Aagaard L, Rossi JJ: **RNAi therapeutics: principles, prospects and challenges.** *Adv Drug Deliv Rev* 2007, **59:**75-86.

64. Grimm D, Streetz KL, Jopling CL, Storm TA, Pandey K, Davis CR, Marion P, Salazar F, Kay MA: **Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways.** *Nature* 2006, **441:**537-541.

65. Li MJ, Kim J, Li S, Zaia J, Yee JK, Anderson J, Akkina R, Rossi JJ: **Long-term inhibition of HIV-1 infection in primary hematopoietic cells by lentiviral vector delivery of a triple combination of anti-HIV shRNA, anti-CCR5 ribozyme, and a nucleolar-localizing TAR decoy.** *Mol Ther* 2005, **12:**900-909.

66. Weinberg MS, Rossi JJ: **Comparative single-turnover kinetic analyses of trans-cleaving hammerhead ribozymes with naturally derived non-conserved sequence motifs.** *FEBS Lett* 2005, **579:**1619-1624.

67. Weinberg MS, Ely A, Passman M, Mufamadi SM, Arbuthnot P: **Effective anti-hepatitis B virus hammerhead ribozymes derived from multimeric precursors.** *Oligonucleotides* 2007, **17:**104-112.

68. Burke DH, Greathouse ST: **Low-magnesium, trans-cleavage activity by type III, tertiary stabilized hammerhead ribozymes with stem 1 discontinuities.** *BMC Biochem* 2005, **6:**14.

69. Saksmerprome V, Roychowdhury-Saha M, Jayasena S, Khvorova A, Burke DH: **Artificial tertiary motifs stabilize trans-cleaving hammerhead ribozymes under conditions of submillimolar divalent ions and high temperatures.** *Rna* 2004, **10:**1916-1924.

70. Seger R, Krebs EG: **The MAPK signaling cascade.** *Faseb J* 1995, **9:**726-735.

71. Qi M, Elion EA: **MAP kinase pathways.** *J Cell Sci* 2005, **118:**3569-3572.

72. Dohlman HG, Thorner JW: **Regulation of G protein-initiated signal transduction in yeast: paradigms and principles.** *Annu Rev Biochem* 2001, **70:**703-754.

73. McCormick F: **Signalling networks that cause cancer.** *Trends Cell Biol* 1999, **9:**M53-56.

74. Hanahan D, Weinberg RA: **The hallmarks of cancer.** *Cell* 2000, **100:**57-70.

75. Pierce KL, Premont RT, Lefkowitz RJ: **Seven-transmembrane receptors.** *Nat Rev Mol Cell Biol* 2002, **3:**639-650.

76. Liu L, Cao Y, Chen C, Zhang X, McNabola A, Wilkie D, Wilhelm S, Lynch M, Carter C: **Sorafenib blocks the RAF/MEK/ERK pathway, inhibits tumor**

angiogenesis, and induces tumor cell apoptosis in hepatocellular carcinoma model PLC/PRF/5. *Cancer Res* 2006, **66:**11851-11858.

77. Shapiro P: **Discovering new MAP kinase inhibitors.** *Chem Biol* 2006, **13:**807-809.

78. Dominguez C, Powers DA, Tamayo N: **p38 MAP kinase inhibitors: many are made, but few are chosen.** *Curr Opin Drug Discov Devel* 2005, **8:**421-430.

79. Dambach DM: **Potential adverse effects associated with inhibition of p38alpha/beta MAP kinases.** *Curr Top Med Chem* 2005, **5:**929-939.

80. Elion EA: **The Ste5p scaffold.** *J Cell Sci* 2001, **114:**3967-3978.

81. Park SH, Zarrinpar A, Lim WA: **Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms.** *Science* 2003, **299:**1061-1064.

82. Elion EA: **Pheromone response, mating and cell biology.** *Curr Opin Microbiol* 2000, **3:**573-581.

83. Chapman S, Asthagiri AR: **Resistance to signal activation governs design features of the MAP kinase signaling module.** *Biotechnol Bioeng* 2004, **85:**311-322.

84. Levchenko A, Bruck J, Sternberg PW: **Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties.** *Proc Natl Acad Sci U S A* 2000, **97:**5818-5823.

85. Wang X, Hao N, Dohlman HG, Elston TC: **Bistability, stochasticity, and oscillations in the mitogen-activated protein kinase cascade.** *Biophys J* 2006, **90:**1961-1978.

86. Esch RK, Errede B: **Pheromone induction promotes Ste11 degradation through a MAPK feedback and ubiquitin-dependent mechanism.** *Proc Natl Acad Sci U S A* 2002, **99:**9160-9165.

87. Esch RK, Wang Y, Errede B: **Pheromone-induced degradation of Ste12 contributes to signal attenuation and the specificity of developmental fate.** *Eukaryot Cell* 2006, **5:**2147-2160.

88. Maleri S, Ge Q, Hackett EA, Wang Y, Dohlman HG, Errede B: **Persistent activation by constitutive Ste7 promotes Kss1-mediated invasive growth but fails to support Fus3-dependent mating in yeast.** *Mol Cell Biol* 2004, **24:**9221-9238.

89. Ingolia NT, Murray AW: **Positive-feedback loops as a flexible biological module.** *Curr Biol* 2007, **17:**668-677.

90. Santos SD, Verveer PJ, Bastiaens PI: **Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate.** *Nat Cell Biol* 2007, **9:**324-330.

91. Kholodenko BN: **Untangling the signalling wires.** *Nat Cell Biol* 2007, **9:**247-249.

92. Bashor CJ, Helman NC, Yan S, Lim WA: **Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics.** *Science* 2008, **319:**1539-1543.

93. Ajo-Franklin CM, Drubin DA, Eskin JA, Gee EP, Landgraf D, Phillips I, Silver PA: **Rational design of memory in eukaryotic cells.** *Genes Dev* 2007, **21:**2271-2276.

94. Kueh HY, Rothenberg EV: **Regulatory gene network circuits underlying T cell development from multipotent progenitors.** *Wiley Interdiscip Rev Syst Biol Med* 2011.

95. Thomson M, Liu SJ, Zou LN, Smith Z, Meissner A, Ramanathan S: **Pluripotency factors in embryonic stem cells regulate differentiation into germ layers.** *Cell* 2011, **145:**875-889.

Chapter 2

Identifying pathway regulators and constructing RNA-based regulatory systems to control decision-making in the yeast mating pathway

Abstract

Synthetic biology is advancing capabilities in constructing genetic circuits that can program biological functions. However, the application of these circuits to the regulation of cell fate has been limited by the means to connect synthetic circuitry to the pathways that govern cellular decision-making. Identifying and controlling the expression of key pathway regulators represents one possibility for connecting synthetic circuits to the processes that govern cell fate. Here, we show that in a model MAPK pathway, the *Saccharomyces cerevisiae* mating pathway, there exist titratable positive and negative regulators of pathway activity and a narrow range of expression of these regulators over which cellular fate diverges. Synthetic circuits designed to control the levels of these regulators were able to divert cells to alternative fates even in the face of antagonistic external signals. By layering these circuits with small-molecule-responsive, RNA-based controllers, we constructed molecular network diverters that specify alternative cell fate decisions in response to exogenous environmental signals. The diverters are designed to remain quiescent in the absence of the environmental input, but activate in the presence of the input signal to route cells to one of two alternative mating fates: “chaste” or “promiscuous”. In wild type cells, pheromone stimulates the mating pathway, inducing cell cycle arrest. Cells exhibiting the chaste phenotype are insensitive to mating cues and do not arrest, whereas promiscuous cells exhibit upregulated signaling through the mating pathway and arrest in the absence of the canonical mating stimulus. Diverters that incorporate a positive feedback architecture enhanced switching to the promiscuous fate, whereas diverters incorporating negative feedback exhibited reduced population heterogeneity. The molecular network diverters control fate decisions without modifying

the host's native genetic material, obviating the need for laborious and potentially infeasible genetic manipulation. The modularity of this control scheme makes this technique broadly applicable to an extensive range of native networks such as those that govern stem cell fate and the establishment and proliferation of cancer.

Introduction

To orchestrate complex, coordinated, multicellular tasks, organisms dynamically program their extracellular space with distributed molecular signals that are processed by individual cells into concerted responses. These extracellular signals activate signaling cascades that induce specific network topologies, leading cells to divergent cellular fates. MAPK cascades are a class of highly conserved signaling pathways that control such key cellular processes as differentiation, mitosis, and apoptosis [1]. Many diseases, including one third of human cancers, result from aberrant signaling through MAPK pathways [2, 3]. In the model eukaryotic organism *Saccharomyces cerevisiae*, multiple MAPK cascades direct cellular fate via divergent regulatory programs. Decisions to upregulate gene expression, halt cell cycle, and change cell morphology are programmed through these pathways as the rational response to changing environmental signals.

The ability to dynamically modulate signaling through MAPK pathways in the face of antagonistic environmental stimuli represents an enormous hurdle in the reprogramming of cell fate decisions. Many of the current strategies for programming cell fate are based on strict control over the extracellular environment [4, 5], presenting limitations in clinical settings and necessitating the development of strategies that are more amenable to therapeutic contexts. Several modeling efforts have posited that levels of signaling molecules are responsible for divergent cell fates [6-8]. Further, experimental studies comparing wild-type response to overexpression and knockouts of signaling pathway components demonstrate that differential pathway activity and cell fate can be achieved via altering the expression levels of particular proteins [9-11]. However, these examples represent extreme perturbations to pathway response. We postulate that

within the continuum of regulator expression levels between wild type and high levels of ectopic overexpression there exists a transitory region over which cell fate diverges (Figure 2.1). In such a model, at subtransition levels of regulator expression, cells mirror the wild-type response. At expression levels above the requisite transitory range, cells adopt an alternative cell fate. By constructing synthetic circuits that can toggle between subtransition and supertransition levels of regulator expression, cellular behavior can be synthetically switched from wild type to programmed alternative fates.

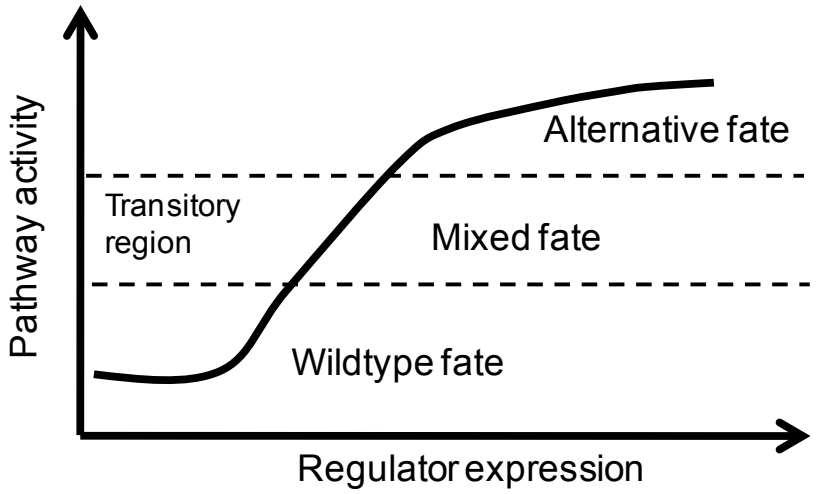


Figure 2.1. Regulator expression modulates pathway activity and over a narrow range of expression transitions to an alternative fate

We present a method for identifying and interfacing with control points in the *S. cerevisiae* mating and osmolarity response pathways. Using a galactose-responsive titration system, we successfully identified both a positive and a negative regulator in the mating pathway and two coexpressed positive regulators in the osmolarity pathway. The regulators of the mating pathway were shown to route cells to alternative fates above a particular threshold of expression. By incorporating these regulators into various network architectures with RNA-based controllers [12], we successfully constructed several types

of molecular network diverters in the mating pathway which conditionally route cells to one of two alternative cell fates based on the presence of distinct environmental signals. Our results demonstrate the utility of conditionally activated control systems for applications modulating cell grow and/or viability as well as the rational tuning of these synthetic circuits via the exchange of well-characterized parts. Further, we show that positive feedback reshapes the pathway activity response curve, shifting the threshold at which cell fate diverges to lower levels of the positive regulator. Finally, we have elucidated design principles and methodology for constructing molecular network diverters that can be readily extended to new pathways and applications.

Results

Identifying key pathway regulators that allow routing of cell fate decisions to alternative phenotypes in the mating pathway

S. cerevisiae activates a three-tiered MAPK cascade upon stimulation with pheromone that upregulates transcription of mating genes, induces cell cycle arrest, and initiates polarized cell growth characteristic of the mating response [8-9] (Figure 2.2). To identify regulators of pathway activity, we individually titrated the levels of signaling proteins in the pathway via an engineered galactose-inducible promoter system (*pGAL1*), which allows for linear, homogeneous regulation of a target gene [13] (Supplementary Figure 2.1). The engineered strain allowed us to identify key pathway regulators by interrogating the effect of titrating mean protein expression levels on pathway activity. We used these strains to determine if overexpression of the regulators was sufficient to cross the transitory range to mediate routing to an alternative fate. Pathway activity was

measured by monitoring expression from a transcriptional fusion construct (*pFUS1*-GFP, GFP fused to a mating responsive promoter), where GFP levels represent a measure of pathway activation, and by observing mating-associated cell cycle arrest via halo assays (see Materials and methods).

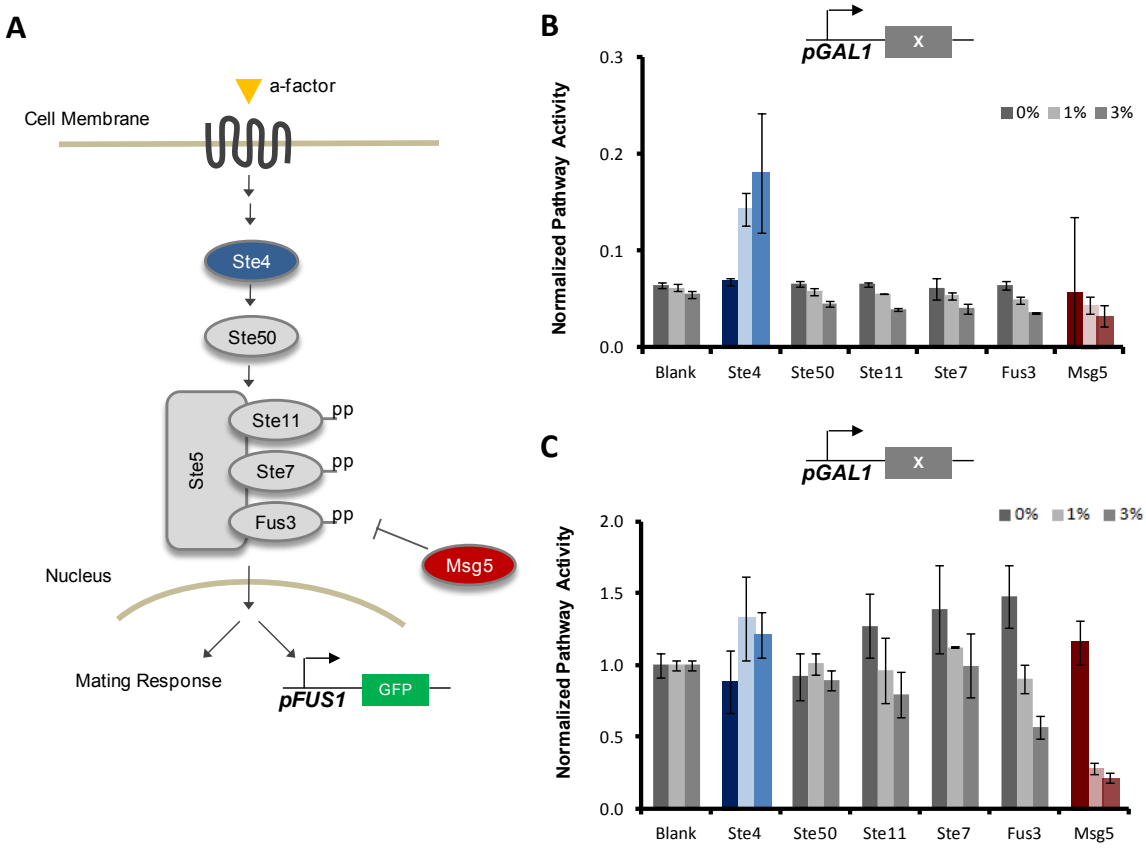


Figure 2.2. Identifying titratable regulators of pathway activity in the yeast mating pathway. A. A molecular view of signaling in the yeast mating pathway. Pheromone (α-factor) binds to a transmembrane receptor to initiate signaling. That binding event is transduced across the cell membrane and relayed down to the G-proteins including, Ste4 and Ste50, an adaptor protein, that in turn relay signaling to the Ste5 scaffold-bound three-tiered MAPK cascade. Phosphorylation of the MAPK Fus3 results in its translocation to the nucleus and is antagonized by the phosphatase Msg5. Fus3 translocation activates a host of transcription factors which ultimately upregulate expression at various mating gene including FUS1. Pathway activity can be monitored via an integrated pFUS1-GFP promoter fusion. Additionally, the mating response is characterized by pheromone-induced cell cycle arrest and polarized cell growth called shmoo formation. **B.** To find pathway activators, cells were monitored in the absence of stimulating pheromone. Ste4 increased pathway activity in a galactose-dependent manner indicating it is a titratable regulator. **C.** To find pathway inhibitors, the pathway was stimulated with pheromone. Msg5 overexpression reduced stimulated pathway activity. Cells were cultured in 0, 0.25 and 1% galactose for 6 hours. Pheromone was added 3 hours post-dilution to stimulate the pathway. Pathway activity was monitored via flow cytometry measurement of cellular fluorescence of GFP from the promoter fusion pFUS1-GFP.

Only two of the six pathway proteins examined in this system exhibit significant effects as positive or negative regulators (Figure 2.2B, C). Overexpression of Msg5 from the engineered galactose-titratable promoter system results in a decrease in pathway activation, as measured from the *pFUS1*-GFP reporter construct, and over a threshold level Msg5 eliminates pheromone-induced cell cycle arrest, as demonstrated by halo assays (Figure 2.3A, B). Conversely, Ste4 overexpression increases pathway activation and reduces cell growth (Figure 2.3D, E). For both regulators, phosphorylated levels of the MAPK Fus3 correlated with GFP levels and cell fate, indicating that the synthetic circuits act through the canonical signaling pathway (Figure 2.3C, F). While previous work supports the role of Msg5 and Ste4 as negative and positive regulators of pathway activity [10, 14], respectively, we have demonstrated that these regulators modulate pathway activity in a dose-dependent manner and above a particular threshold of expression each regulator directs cell fate away from the wild-type response to an alternative phenotype. High levels of Msg5 inhibit pathway activation in the presence of pheromone, routing cells to a "chaste" fate. Above a threshold level of Ste4, cells constitutively activate the pathway in the absence pheromone to adopt a "promiscuous" fate. Our identification of only two pathway regulators from this screen supports the hypothesis that within many signaling pathways regulatory architectures filter out perturbations in the expression of signaling components [15, 16]. However, despite the native control schemes, we identified two control points where overexpression of the regulatory proteins dictates the level of pathway activity.

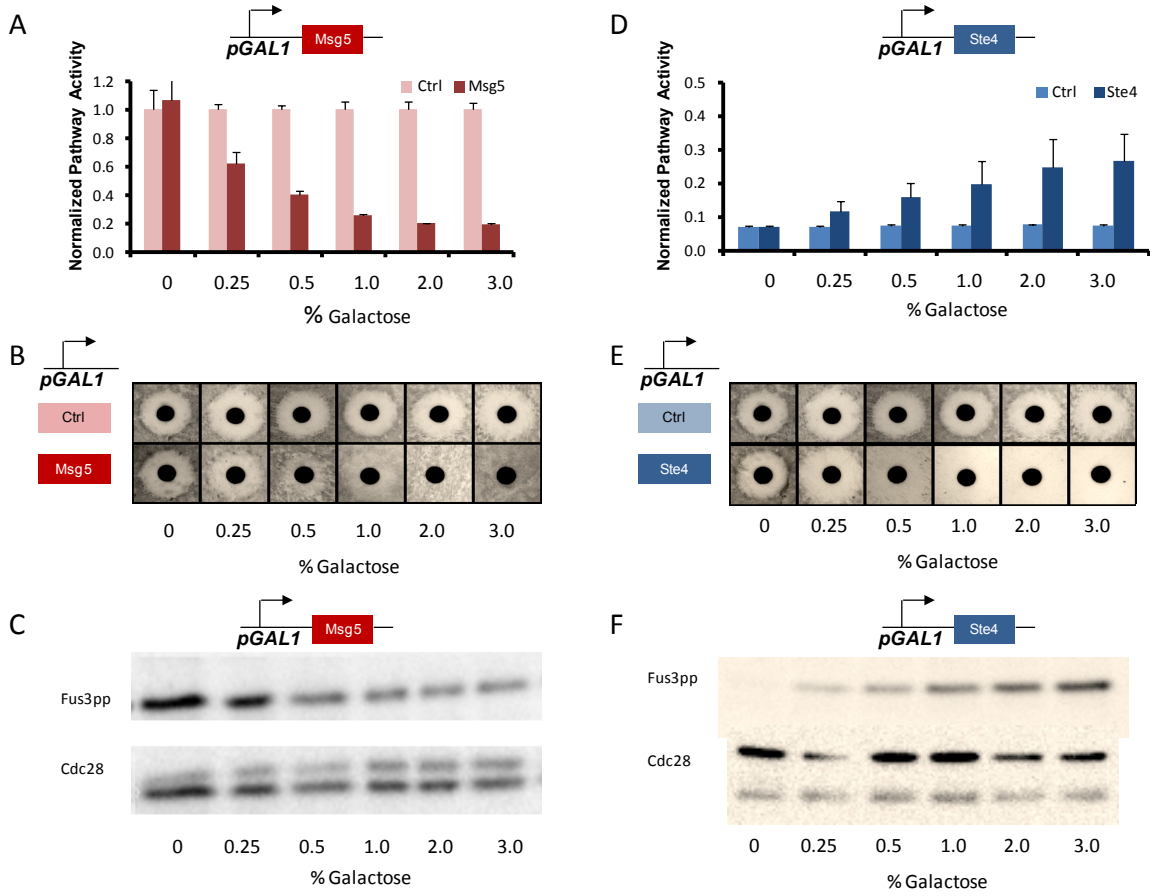


Figure 2.3. Overexpression of Msg5 and Ste4 modulates pathway activity and routes cells to an alternative fate. A. Increasing Msg5 overexpression results in reduced pathway activity over the galactose range as measure by pFUS1-GFP levels **B.** Halo assays at various concentrations of galactose indicate that as galactose concentration increases Msg5 overexpression inhibits pheromone-induced cell cycle arrest leading to a "chaste" phenotype. **C.** Phosphorylated levels of Fus3 correspondingly drop across the galactose range while the Cdc28 control levels remain constant. **D.** For Ste4 overexpression, pathway activity increases as galactose levels increase. **E.** Halo assays at various concentrations of galactose indicate that as galactose concentration increases Ste4 overexpression generates pheromone-independent cell cycle arrest or "promiscuous" fate. **F.** Phosphorylated levels of Fus3 correspondingly drop across the galactose range while the Cdc28 control levels remain constant.

Identifying pathway regulators in the osmolarity pathway

To examine the flexibility of the synthetic titration system for identifying pathway regulators, we applied it to a second yeast MAPK pathway, the osmolarity response pathway. The osmolarity response is triggered by high osmotic pressure that initiates signaling at two different receptors Sho1 and Sln1, which converge on a three-tiered

MAPK cascade (Figure 2.4A). Signaling results in phosphorylation and translocation of the dedicated osmo-MAPK Hog1 to the nucleus, and ultimately upregulation of osmo-genes and production of glycerol. Activation of the osmolarity response is also known to repress the mating pathway. We used the engineered galactose-inducible strain to evaluate the potential of several signaling proteins as titratable positive pathway regulators. Negative regulators were not explored with this method as inhibiting the osmolarity pathway in yeast challenged with high osmotic pressure results in cell death [17].

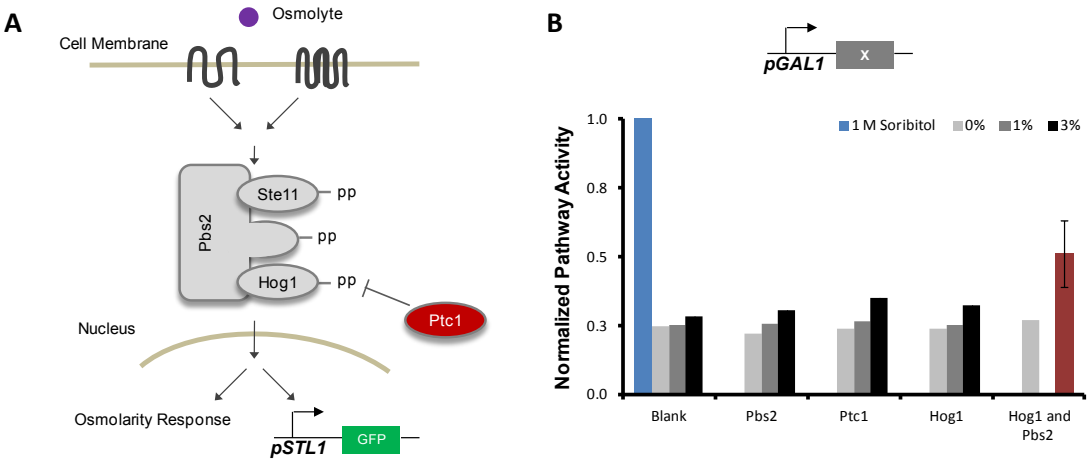


Figure 2.4. Identifying titratable regulators in the yeast osmolarity pathway. A. A molecular view of signaling in the yeast osmolarity pathway. High osmotic pressure initiates signaling via two transmembrane sensor proteins. Signaling is relayed to the Pbs2 scaffold-bound three-tiered MAPK cascade. Phosphorylation of the MAPK Hog1 results in its translocation to the nucleus and is antagonized by the phosphatase Ptc1. Hog1 translocation activates a host of transcription factors which ultimately upregulate expression at various osmo genes including STL1. Pathway activity can be monitored via an integrated pSTL1-GFP promoter fusion. Additionally, the osmolarity response is characterized by glycerol production. **B.** Increasing pathway activity in the osmolarity pathway requires overexpression of both Hog1 and Pbs2. To find pathway activators, cells were monitored in the absence of sorbitol. Cells were cultured in 0, 1, and 3% galactose for 6 hours. Pathway activity was monitored via flow cytometry measurement of cellular fluorescence of GFP from the promoter fusion pSTL1-GFP.

The impact of overexpression of three pathway proteins on osmolarity pathway activation was examined using the galactose-titratable expression system. Pathway activation was measured via a transcriptional fusion between a promoter that is activated through osmolarity pathway stimulation and GFP (*pSTL1*-GFP). Even at high induction

levels, the overexpression of single pathway proteins resulted in very small increases in pathway activation (Figure 2.4B). We hypothesized that co-overexpressing both the scaffold-kinase, Pbs2, and the MAPK Hog1 would lead to more substantial increases in pathway activity. We observed that co-overexpression of Pbs2 and Hog1 significantly increased pathway activation over single component expression.

Evaluating whether an alternative phenotypic fate can be achieved by controlling activation of the osmolarity pathway via regulator overexpression encountered several challenges. First, unlike the mating pathway (e.g., cell cycle arrest) the osmolarity pathway does not have a well-defined phenotypic readout. As a proxy, Hog1 translocation to the nucleus has been used as a measure of pathway activation using a strain composed of a fluorescently labeled Hog1 and nuclear protein marker [18]. In the native pathway, Hog1 translocation occurs rapidly following hyperosmotic shock and pathway activity attenuates within minutes, returning nuclear Hog1 levels to pre-stimulated levels. This timing dependence poses an issue in determining when to assay Hog1 translocation for our system in which the pathway is not stimulated by osmotic shock but instead by overexpression of protein components established over the course of the 6 hour assay. Similar issues are associated with evaluating phosphorylated Hog1 as a measure of pathway activity since the phosphorylation time scale mirrors that of translocation. Due to the challenges associated with determining fate from pathway activation, we did not perform this evaluation for the engineered systems. With significant method development it may be possible to evaluate fate routing in the osmolarity pathway. However, in light of the time scale of action for the osmolarity response, transcriptional feedback control may not be aptly suited to controlling the

osmolarity pathway. Under such time scale constraints, transcriptional control systems may modulate the initial conditions that guide the response, but dynamic modulation of signaling via feedback would be necessarily precluded.

Building molecular network diverters for programming cell fate decisions

A molecular network has a particular topology which shapes the flow of signal in the pathway. Reshaping the network in decision-making pathways, such as MAPK pathways, offers the potential to divert the endogenous response to an alternative fate. Network reshaping requires identifying and plugging into a network regulatory node, a sensitive control point in the pathway for directing pathway activity and fate. Pathway regulators govern the function of their cognate regulatory nodes. Changes in the native network shape can be induced by raising the profile of a pathway regulator. A molecular network diverter conditionally routes cells to an alternative fate by transducing signals in the environment into changes in network shape via the modulation of the pathway regulator profile (Figure 2.5). A molecular network is composed of a switch which encodes a sensing function that recognizes the environmental signal. Increasing the input concentration increases the signal transduced to the actuator element within the switch (Figure 2.5A). The activity of the actuator combined with the strength of the modulator dictates the profile of the pathway regulator. As the concentration of the environmental input increases, the profile of the imbedded pathway regulator rises. Once the profile exceeds a particular threshold, the regulator docks to the network regulatory node and induces changes in the native molecular network, diverting the pathway response (Figure 2.5B).

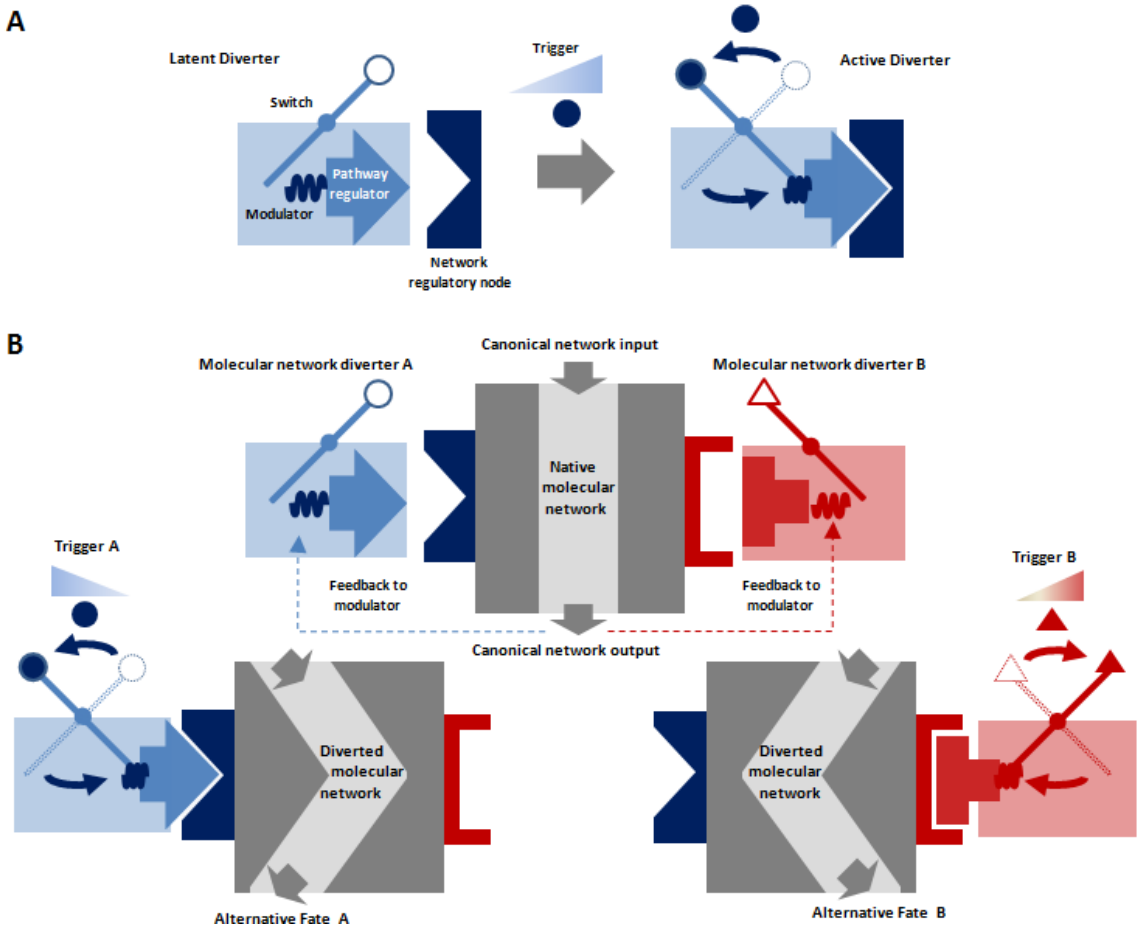


Figure 2.5. Reshaping the native molecular network with molecular network diverters. A. The latent diverter is activated by increasing trigger concentrations with flip the switch which raises the profile of the pathway regulator via the modulator. Above a certain threshold, the pathway regulator docks to the regulatory node to reshape the native network. **B.** Trigger binding to its cognate sensor flips the switch. Binding is transduced via the switch actuator to the modulator and finally to the imbedded pathway regulator, raising the profile of the regulator which plugs into its cognate regulatory node, reshaping the molecular network and diverting fate. The diverter function can be tuned by the properties of the switch, the modulator, and addition of feedback to the modulator.

The properties of the molecular network diverter can be tuned via the genetic parts selected to compose the diverter. The diverter's component functions can be mapped to particular genetic parts (Figure 2.6). The modulator maps to a promoter, where the promoter strength determines how strongly changes in the activity of the switch actuator are translated into changes in the profile of the pathway regulator. For the mating pathway, the pathway regulator is a gene cassette encoding either Ste4 or Msg5. The switch maps to RNA-based controllers.

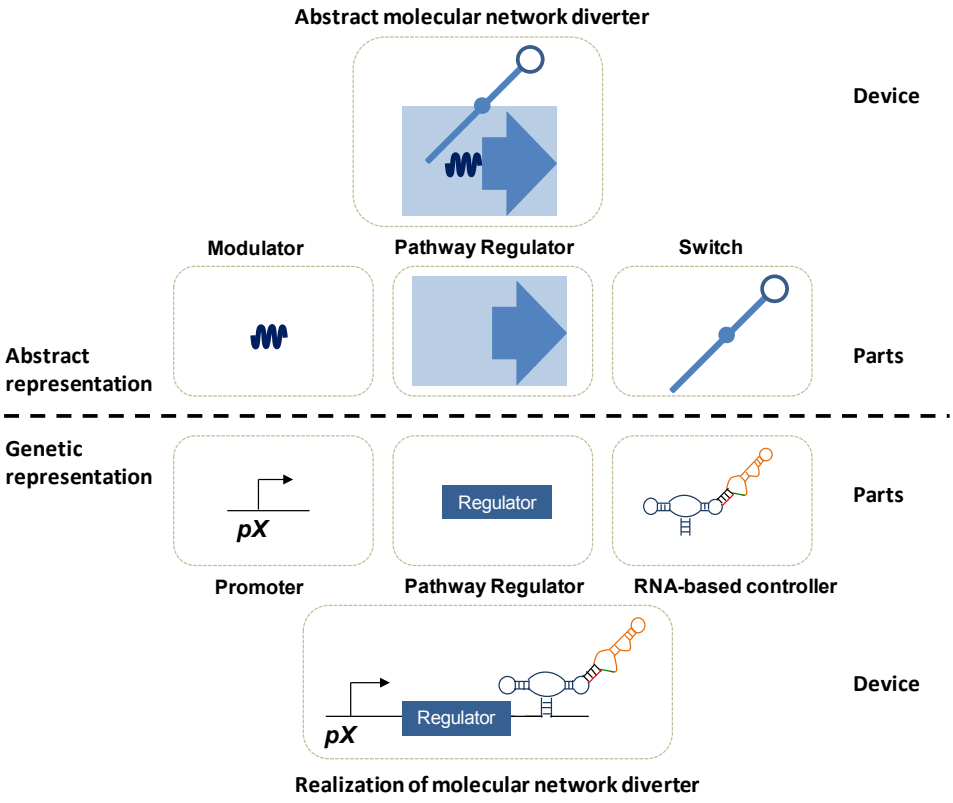


Figure 2.6. Implementation of a molecular network diverter from various genetic parts. The molecular network diverter is composed with a modulator, pathway regulator, and a switch into the larger device. Mapping these parts to their genetic representation, the modulator is a promoter, the pathway regulator is a gene that regulates pathway activity, and the switch is an RNA-based controller. Assembling the genetic parts in the proper configuration allows the realized molecular network diverter to conditionally regulate pathway activity and divert cell fate.

Construction of a molecular network diverter requires the selection of regulatory elements with the appropriate strength and range for modulating the pathway regulator levels to conditionally route cellular fate. The selection of the promoter driving the expression of the regulator is a key component in setting the activity of the molecular network diverter. Constitutive promoters of varying strength achieve different levels of regulator expression (Figure 2.7). The pathway activity response curve indicates where a promoter-regulator pair falls on this curve, which dictates whether cells route to the wild-type or alternative fate. Using a promoter responsive to pathway activation results in the construction of a feedback molecular network diverter. Feedback in molecular networks

can be used to change the shape of the pathway activity response curve. Negative feedback has been shown to antagonize ultrasensitivity, converting sigmoidal responses to linear responses [19]. Positive feedback architectures can be used to establish ultrasensitivity, reducing the threshold of expression required to trigger switching to an alternative fate [20]. The strength of positive feedback represents a sensitive control point, which dictates whether cells adopt the wild-type or alternative fate. We used four constitutive promoters of varying strength and a feedback promoter to construct and evaluate molecular network diverters (Supplementary Figure 2.2).

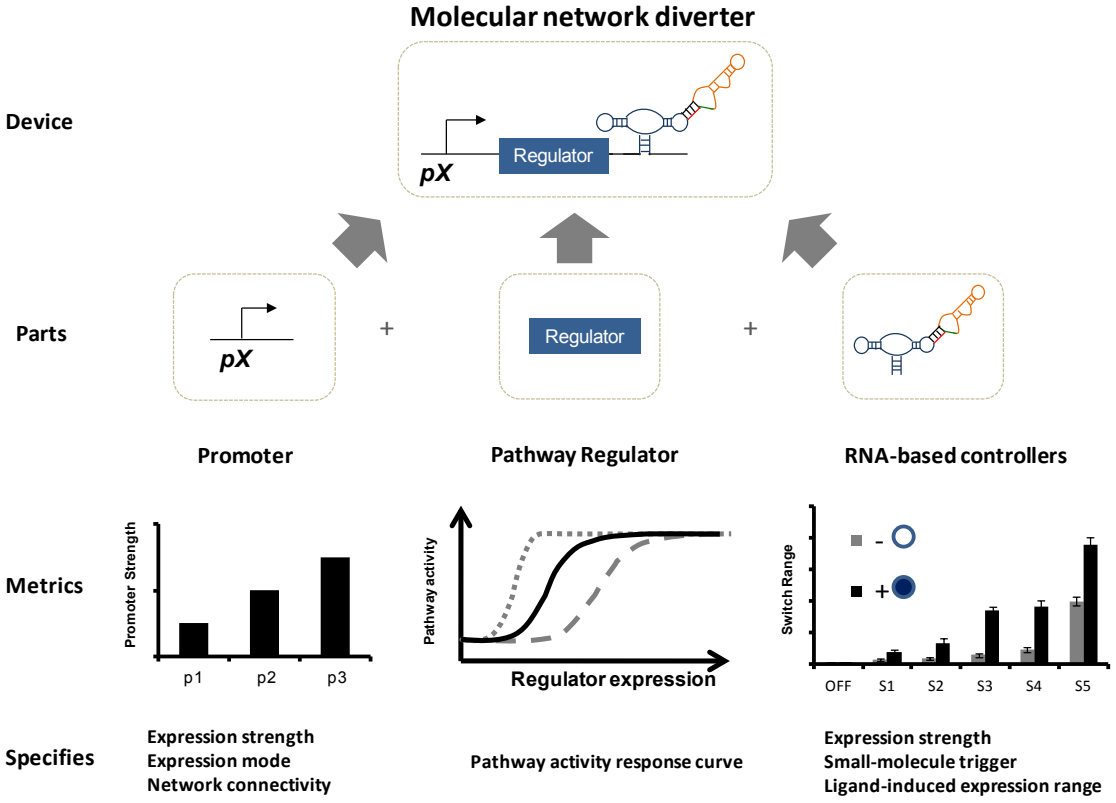


Figure 2.7. Composition of a molecular network diverter from well-defined parts. The molecular diverter is composed of a promoter, a pathway regulator, and a RNA-based controller, also called a switch. The promoter specifies the expression mode (e.g., constitutive or feedback) which determines the network connectivity. The promoter and switches combine to determine the expression strength and thus the activity of the diverter. The pathway regulator's interaction with the native molecular network determines the pathway response curve. The pathway response curve can be altered by changes in network connectivity due to feedback expression of regulators. Switches specify which small-molecule input regulates expression and the range of expression across ligand concentrations.

RNA controllers provide a second layer of control by which to tune pathway regulator levels and allow the molecular network diverter to respond to environmental signals. Small-molecule-responsive RNA-based controllers, also called RNA switches, are an engineered class of noncoding RNA that allows for conditional control from any promoter-gene pair via a chemical trigger [12]. The RNA switches act post-transcriptionally to destabilize transcripts in an input-dependent manner, thereby upregulating or downregulating target protein levels in response to increasing concentrations of environmental triggers. The regulated output achieved by these programmable RNA controllers spans a wide window of dynamic ranges and regulatory stringencies (Supplementary Figure 2.3). The switches used in this study respond to one of two chemical triggers, allowing for two independent channels through which to simultaneously regulate Msg5 and Ste4 expression. The relative switch regulatory activity is independent of other circuit components (e.g., gene, promoter) such that the activity relative to the ON state control (nonswitching) is expected to be consistent across different promoter-gene pairs, supporting rational programming of the molecular network diverter properties. The ON and OFF state controls (nonswitching) demark the highest and lowest potential expression levels, respectively, from a promoter-gene pair coupled to the RNA-based controllers. Since the ON and OFF state expression levels of the RNA switches fall within these extrema, differential fates from the controls must be observed in order to construct molecular network diverters that alter cell fate decisions (Figure 2.8).

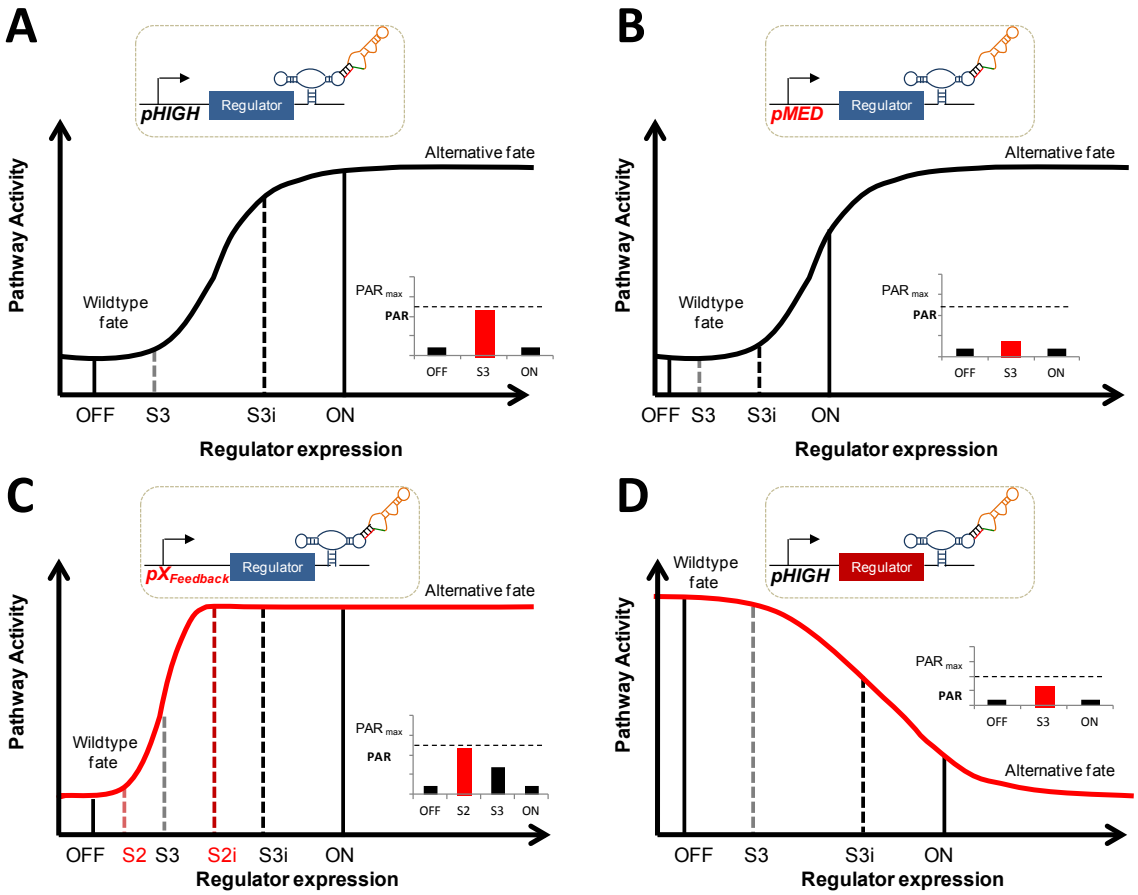


Figure 2.8. Optimal configuration of the molecular network diverter requires tuning via the selection of components with the requisite metrics. A. An optimally configured diverter spans the steep linear range of the pathway response curve to yield a PAR value approaching PAR $_{max}$. Switch S3 in the absence of input maintains the wild-type fate with low pathway activity (grey dashed line). With the addition of input (S3i), S3 switches to high pathway activity routing to the programmed alternative fate (black dashed line). **B.** Reducing the promoter strength results in a suboptimally configured diverter. **C.** Implementing feedback expression of a positive regulator introduces ultrasensitivity. As a result, a more stringent switch (with lower basal levels) is required to achieve optimal diverter performance. **D.** Substituting a negative regulator for a positive regulator enables construction of a negative diverter. The negative diverter attenuates activity in the stimulated pathway. PAR (pathway activation ratio) is the ratio of the pathway activity in the presence and absence of the input for positive diverters. For negative diverters, PAR is the inverse of this ratio.

A molecular network diverter requires that regulator expression spans the transitory range such that expression is above the requisite threshold in the presence of the input signal, while sub-transition levels are maintained in the absence of the signal. Optimal performance of the molecular network diverter requires tuning via the selection of components with the requisite metrics (Figure 2.8). In an optimally configured

diverter, the switch maintains wild-type behavior in the absence of environmental input and routes to the alternative fate in the presence of the input (Figure 2.8A). The efficiency of the molecular network diverter's response to the input signal is measured by the pathway activation ratio (PAR), the ratio of the pathway activity in the presence and absence of the input for positive diverters. For negative diverters, PAR is the inverse of this ratio. An optimally configured diverter spans the steep linear region of the pathway response curve, resulting in a PAR value approaching PAR_{max}. For positive diverters, PAR_{max} is the ratio of the maximum to minima of the response curve as shown for S3 (Figure 2.8A, inset); for negative diverters, PAR_{max} is the inverse ratio.

The modular components within the molecular network diverter architecture provide an array of tools for tuning the activities of these controllers to best interface with endogenous cellular networks. For example, changing the promoter in the regulator-switch pair can significantly alter the diverter's PAR value (Figure 2.8A, B). Additionally, the introduction of positive feedback into the network diverter architecture can result in shifts to the response curve that introduce ultrasensitivity and ultimately require more stringent RNA controllers to achieve optimal PAR values (Figure 2.8C). As another example, exchanging a positive pathway regulator in a positive network diverter for a negative regulator enables the construction of a negative network diverter in which activity is attenuated (Figure 2.8D).

Engineering a positive molecular network diverter that conditionally routes cells to a "promiscuous" phenotype

To construct a positive molecular network diverter, we paired Ste4 with the four constitutive promoters of varying strength (Supplementary Figure 2.2) and the ON state control and evaluated the pathway response. For the three lower promoter strengths, cells weakly adopt a promiscuous fate characterized by diminished growth relative to cells harboring a control construct that lacked Ste4. Cells transformed with a construct harboring Ste4 under the control of the strongest promoter repeatedly failed to produce colonies. No phenotypic differences emerge across the range of the three lower promoter strengths examined (Supplementary Figure 2.4A). We paired pHIGH-Ste4 with tetracycline-responsive switches and nonswitch controls to determine the pathway response curve to regulator expression and the potential for phenotypic switching from this molecular network diverter architecture. Mild input-triggered phenotypic switching from wild-type to a promiscuous phenotype was observed from switch S4tc (Figure 2.9A). In contrast, no phenotypic switching was observed from any of the diverters in which the lower strength promoter, pLOW, was paired with Ste4 (Supplementary Figure 2.4B). We postulated that feedback expression of Ste4 might yield more robust phenotypic switching by introducing ultrasensitivity into the pathway response curve and by reinforcing pathway activation. The introduction of feedback expression of Ste4 into the network diverter architecture generated strong phenotypic switching from two tetracycline-responsive switches, S3tc and S4tc (Figure 2.9B).

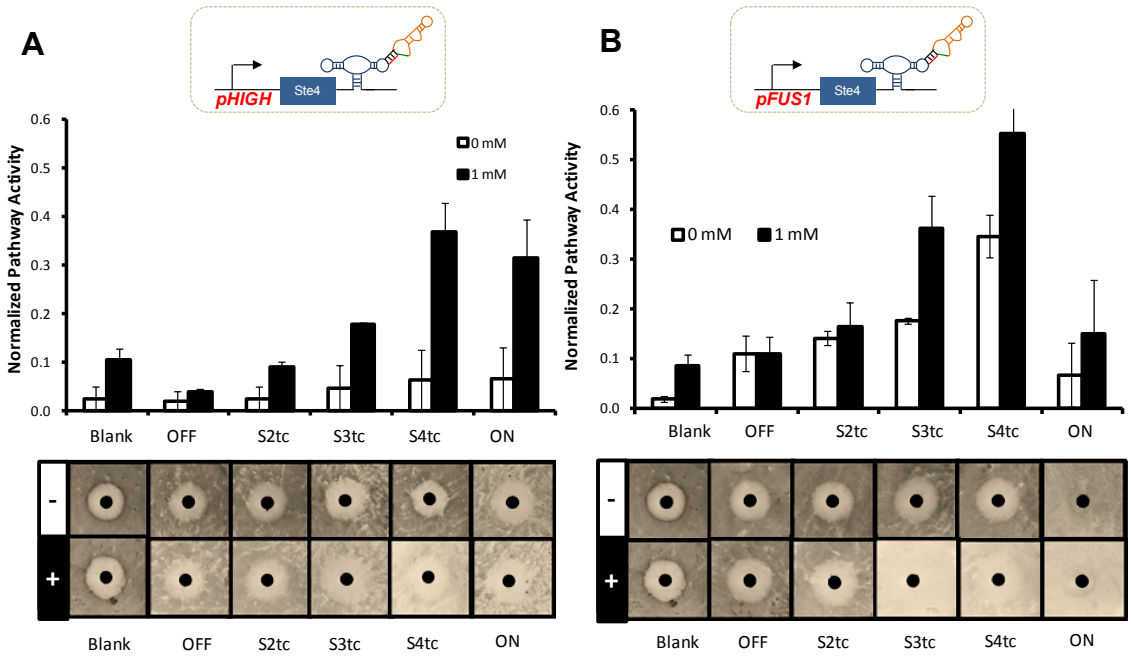


Figure 2.9. Tetracycline-inducible positive network diverters of different architectures conditionally route cells to the promiscuous phenotype. A. Pairing tetracycline-inducible switches of increasing strength with constitutive expression of Ste4 results in increasing levels of pathway activity. Addition of tetracycline increases pathway activity for each switch construct. For S4tc, plate assays in the absence of tetracycline show a wild-type halo, while at 1 mM tetracycline cells adopt the promiscuous phenotype. **B.** Constructing a positive feedback loop amplifies ligand-induce phenotype switching. S3tc and S4tc show wild-type halos in the absence of tetracycline and a strong promiscuous response to tetracycline with nearly undetectable levels of cellular growth.

A second set of diverters were composed from theophylline-responsive switches in both expression modes to demonstrate that routing via the molecular network diverters can be adapted to different target small-molecule inputs. Similar trends were observed between the tetracycline and theophylline-responsive switches. As expected, switches with similar basal levels and ranges achieve similar pathway activities and routing capabilities indicating the routing effect is independent of the specific identity of the input signal (Supplementary Figure 2.5). Further, the data support the ability to rationally tune molecular network diverters by the exchange of modular, well-defined parts.

The data also suggest important tradeoffs between activity and timing in constructing programs that modulate cellular growth. For example, the ON state control

for feedback expression of Ste4 exhibited significant growth despite being programmed for constitutive arrest. We postulated that strong Ste4 overexpression can place selective pressure against cells with highly active pathways, such that over time the population is dominated by cells resistant to pathway activation. Consequently, cells harboring these programs display a low pFUS1-GFP profile and exhibit growth within the expected halo region indicating resistance to pheromone-induced cell cycle arrest. At superthreshold feedback strengths, only cells resistant to activation of the mating program and/or the synthetic positive feedback loop are selected over time. To test this theory, we observed the behavior and GFP profile of cells harboring the feedback molecular network diverters in the presence and absence of the trigger molecule over time. The GFP profiles of cells harboring pFUS1-Ste4-S4tc dropped when grown in the presence of tetracycline for two days in contrast to cells grown in the absence of tetracycline (Supplementary Figure 2.6A) The plasmid marker levels measured via a constitutively expressed mCherry cassette showed normal variation across the samples. Thus, the mCherry levels did not show a correlation with pathway activity, indicating the plasmid is maintained at similar levels in cells grown in the absence or presence of tetracycline (Supplementary Figure 2.6B). Additionally, uniform growth across the plate in halo assays indicates that cells grown at superthreshold levels over an extended time are resistant to pheromone-induced cell cycle arrest (Supplementary Figure 2.6C). Above a threshold level of pFUS1-Ste4 expression only cells with minimal pathway activity escape persistent cell cycle arrest. These data suggest that the ultrasensitivity introduced by the positive feedback architecture lowers the threshold to activation, capturing nearly the entire population in the high activity region of the curve and broadly enforcing a strong selective pressure

against pathway activation. Ultrasensitivity is not introduced into cells programmed with constitutive expression of Ste4, potentially explaining why similarly strong resistance to mating is not observed in cells harboring network diverters encoding constitutive Ste4 expression. These results highlight the utility of programs that run quiescently until activated for applications in which the program necessarily modulates cellular processes such as growth rate and/or viability.

By compiling the pathway activity data for Ste4, we determined the pathway response curve to varying Ste4 levels, identified the transitory range across which cell fate diverges, and examined differences between constitutive and feedback expression of Ste4 observed in the traces of the pathway response curve. For constitutive expression of Ste4, we traced pathway activity as a function of the calculated Ste4 expression and identified the transitory range over which cell fate diverges. The data points show a steady increase in pathway activity with increasing Ste4 expression up to a threshold followed by a precipitous drop (Figure 2.10A). Only cells induced to super-threshold expression levels show high pathway activity as expected from results indicating that over time high Ste4 expression selects for mating resistant cells. The drop in pathway activity coincides with the divergence of fate observed in the halo assays (Figure 2.10B). For feedback expression, the transitory range resides at Ste4 expression levels that are almost four times lower than those for constitutive expression (Figure 2.10C, D). The results support our hypothesis that positive feedback introduces ultrasensitivity to the pathway response curve, shifting the threshold to lower Ste4 expression levels.

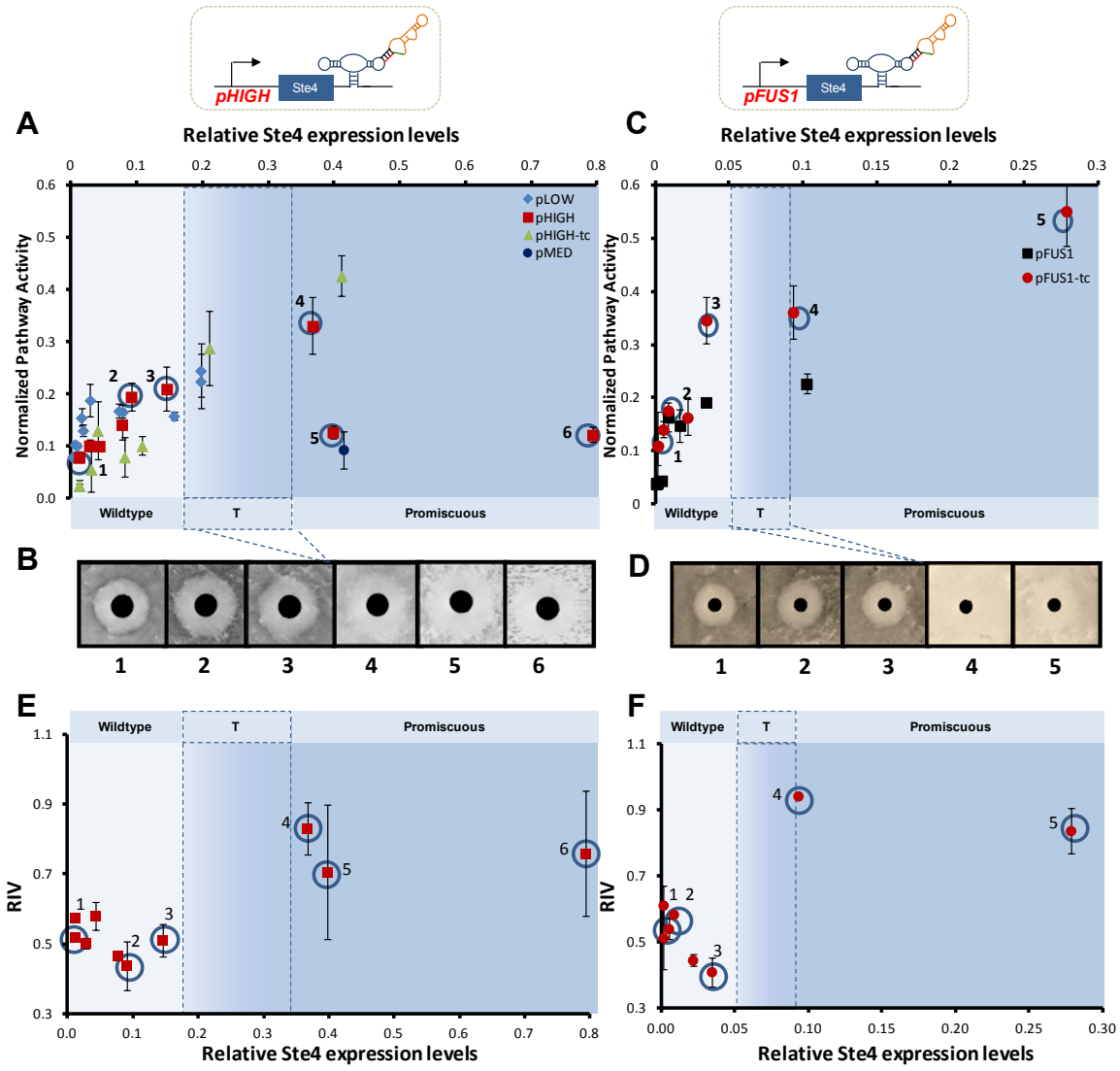


Figure 2.10. Tracing the pathway response curve of Ste4 expression to the promiscuous fate. A. Tracing the pathway response to constitutive Ste4 expression shows a steady increase in pathway activity with increasing relative Ste4 expression up to a threshold followed by a precipitous drop. Pathway activity increases from 2% to 35%. Above 40% expression, pathway activity drops (for calculated Ste4 expression levels see Materials and Methods). **B.** This drop coincides with the divergence of fate shown in the halo assays. Plate data correspond to points with numbered circles in A selected from the pHIGH-SX data set (Supplementary Figure 2.4). **C.** Tracing the pathway response with feedback Ste4 expression shows a narrower transitory range shifted to lower Ste4 expression levels. Pathway activity increases from 0% to 5%. Only samples induced with ligand show Ste4 levels above 10%. **D.** Above 10% Ste4 expression levels, plate assays show cells adopting the promiscuous fate. Plate data correspond to points with numbered circles in A selected from the pFUS1-SXtc data set (Supplementary Figure 2.9B). **E.** Relative intensity values (RIV) are low at low Ste4 expression levels and high at high Ste4 levels, indicating that the difference between inner and outer cell growth as measure by intensity is reduced above a particular level of Ste4 expression (see Methods for calculation of RIV). **F.** RIV values for feedback expression of Ste4 show two regimes of behavior. At low Ste4 expression, RIV values are low, but at high Ste4 expression cells transition to high RIV values indicating cells undergo pheromone-independent cell cycle arrest. "T" indicates transitory region.

We have demonstrated the construction of positive network diverters that route cells to the promiscuous fate. Additionally, we have shown that network diverters that incorporate a positive feedback architecture via a pathway responsive promoter exhibit enhanced fate switching compared to nonfeedback network diverter architectures controlling the same positive regulator, Ste4. Positive feedback loops are known to be important in developmental decision-making processes [21]. Thus, the described feedback molecular network diverters may provide insight into effective strategies for guiding the development of tissues from progenitor cells.

Building a negative molecular network diverter that conditionally routes cells to a "chaste" phenotype

A negative molecular network diverter was constructed by pairing Msg5 with four promoters spanning a range of expression levels. We examined pathway activation in cells harboring these constructs under saturating pheromone conditions. Each construct exhibited routing to the chaste fate (Supplementary Figure 2.7). A set of negative network diverters of varying strength and range that are orthogonal to the positive diverters were constructed by pairing the strong promoter (pHIGH) and the theophylline-responsive switches (Figure 2.11A). Having constructed both tetracycline- and theophylline-responsive positive diverters, negative diverters responsive either small-molecule would provide for an orthogonal pair of diverters each responsive to its particular input signal. While we anticipate that tetracycline-responsive switches would mirror the phenotypic results observed with theophylline, we chose theophylline to regulate the negative diverter to mitigate any potentially confounding issues with pFUS1-

GFP analysis associated with tetracycline autofluorescence. The low expression level of Msg5 associated with the OFF state control allowed pathway activation in the presence of pheromone, leading to high GFP levels, a low PAR value, and wild-type halo formation (Figure 2.11B, C, D). In contrast, the high expression level of Msg5 associated with the ON state control inhibited pheromone-induced pathway activation, thereby eliminating cell cycle arrest and resulting in low GFP levels. Between S2 and S4 cell fate diverges from the wild-type to chaste phenotype. The basal expression levels of Msg5 associated with S1 and S2 in the absence of theophylline are sufficiently low to permit elevated pathway activity in response to pheromone and the characteristic halo formation as is observed for the OFF state control. Both S1 and S2 attenuated pathway activity in the presence of input and showed significant phenotypic switching to the "chaste" fate as is observed for the ON state control.

The molecular network diverters composed with the pHIGH-Msg5 promoter-regulator pair highlight key design considerations associated with tuning network diverters to cross the associated transitory range to alternative phenotypes. The pathway response curve requires switches with stringent control over basal expression levels to maintain quiescence in the absence of the environmental trigger. While switch S4 shows a larger SAR value than S1 and S2 (Supplementary Figure 2.3A), the basal expression level from the diverter incorporating S4 inhibited pathway signaling sufficiently to eliminate cell cycle arrest in the presence of pheromone even in the absence of the input signal, yielding a low PAR value (Figure 2.11). In contrast, switches S1 and S2 exhibit a more stringent control profile than that of S4, and diverters incorporating these switches exhibit larger PAR values and achieve rerouting of cell fate in response to the input

signal. The results highlight the importance of tunable controllers, including those exhibiting stringent profiles, in targeting thresholds of key pathway components to achieve desired phenotypic switching.

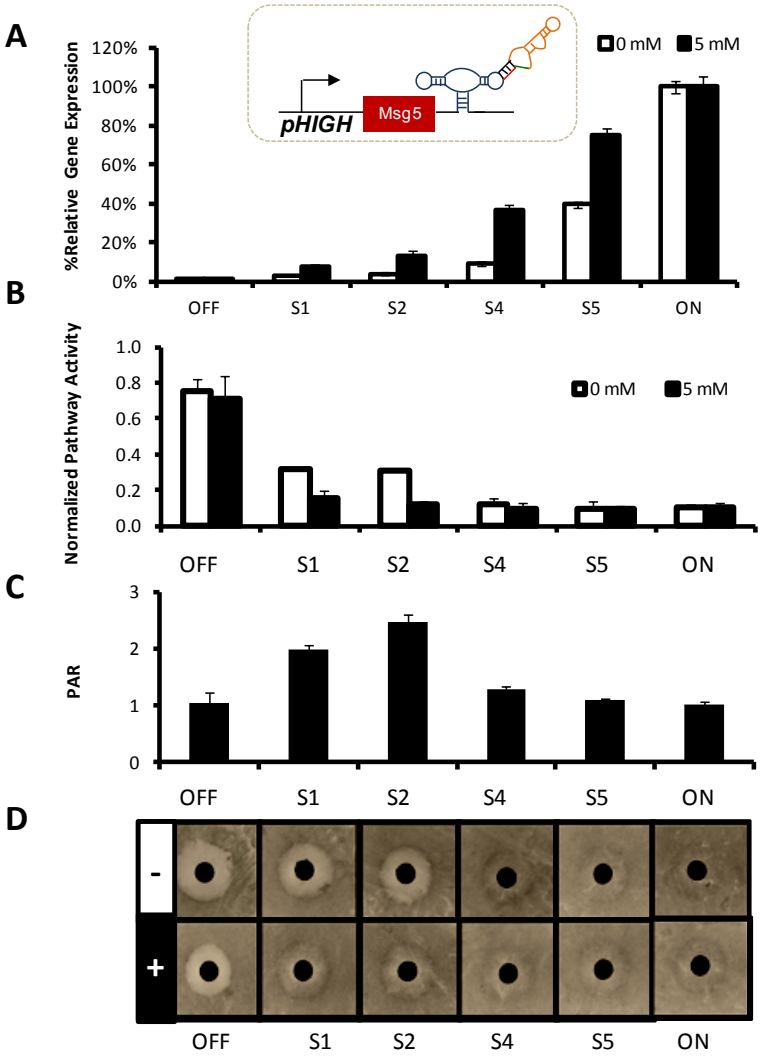


Figure 2.11. Theophylline-inducible molecular network diverters constructed with constitutive Msg5 expression conditionally route cells to the chaste phenotype. A. and **B.** Pairing theophylline-inducible switches of increasing strength from S1 to S4 with constitutive expression of Msg5 results in decreasing levels of pathway activity. **C.** Addition of theophylline reduces pathway activity most significantly for S1 and S2 as indicated by the higher PAR values. **D.** For S1 and S2, plate assays in the absence of tetracycline show a wild-type halo, while at 5 mM theophylline cells adopt the chaste phenotype.

Negative feedback loops have been shown to reduce population heterogeneity [22]. We examined whether improved fate routing could be achieved by integrating a

feedback architecture with the negative molecular network diverters. The negative feedback diverters were composed by pairing the pFUS1-Msg5 cassette with various theophylline-responsive switches (Figure 2.12). The low expression level of Msg5 associated with the feedback diverter incorporating the OFF state control allowed pathway activation in the presence of the pheromone, leading to high GFP levels and wild-type halo formation (Figure 2.12B, D). In contrast, the high expression level of Msg5 associated from the feedback diverter incorporating the ON state control inhibited pheromone-induced pathway activation, thereby eliminating cell cycle arrest and resulting in low GFP levels. Cell fate diverged from the wild-type to chaste phenotype between S2 and S4. The entire set of feedback diverters showed low PAR values, indicating a weak ability to route from wild-type to chaste fate (Figure 2.12C). Weak phenotypic switching from the negative feedback diverters incorporating switches S1 and S2 was confirmed by plate assays (Figure 2.12D). Diverters incorporating negative feedback may linearize the pathway response curve to Msg5 expression. As a result of this linearization, changes in Msg5 expression may be transduced into smaller changes in pathway activity, thus preventing phenotypic switching.

The feedback diverters did not inhibit pathway activation as greatly as the nonfeedback molecular diverters (as measured by transcriptional activation); however, the more active molecular diverters effectively prevented halo formation in the plate assays. Comparing the histograms for the ON and OFF state controls, we observe that while both expression modes reduce the mean of the high GFP population (and thus reduce the level of pathway activation), constitutive expression results in the establishment of a second low GFP population (Supplementary Figure 2.8A, B). GFP

histograms and calculated mean values of the high GFP populations are comparable between the two expression modes (Supplementary Figure 2.8C, D and Supplementary Figure 2.9). These results suggest that cell fate switching at the population level occurs when the mean pathway activity of the high population crosses the critical threshold.

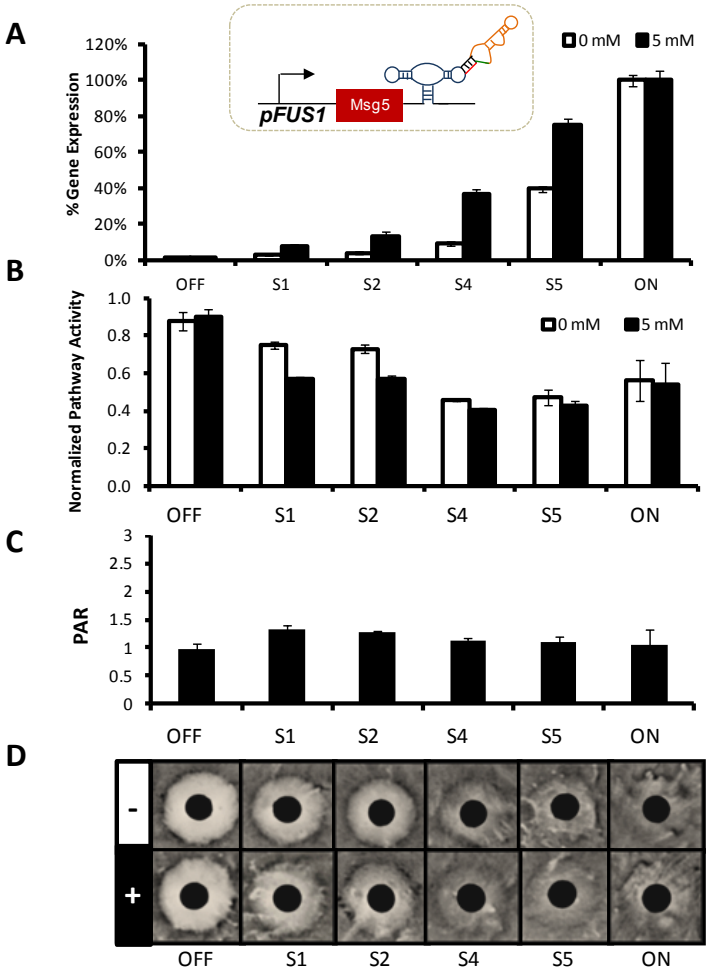


Figure 2.12. Theophylline-inducible molecular network diverters constructed with feedback expression of Msg5 route cells to the chaste phenotype. A. Pairing theophylline-inducible switches of increasing strength from S1 to S5. **B.** Constructing a negative feedback loop with Msg5 results in reduced pathway activity across the increasing switch range. **C.** PAR values are noticeably low compared to constitutive expression of Msg5 (Figure 2.11) **D.** Above S2, cells adopt the chaste phenotype. Only mild phenotypic switching is observed in S1 and S2.

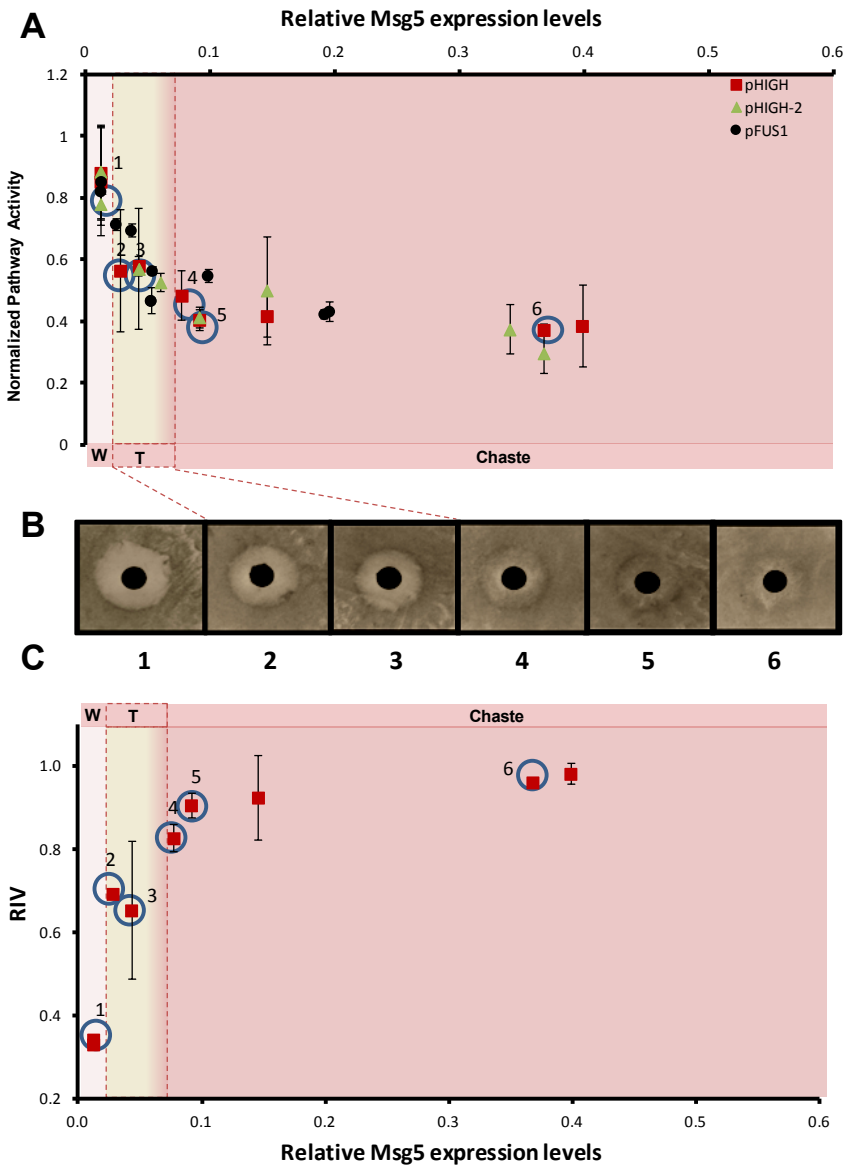


Figure 2.13. Tracing pathway response to Msg5 expression to the chaste fate. A. Pathway activity significantly decreases over a narrow window of relative Msg5 expression (0 – 0.1). **B.** Above 0.08 expression cells adopt the chaste fate. Plate data correspond to points with numbered circles in A selected from the pHIGH-SX data set (Figure 2.11). Relative Msg5 expression was calculated as described previously. **C.** Relative intensity values (RIV) for constitutive expression of Msg5 show two regimes of behavior. At low Msg5 expression (0 – 0.1 relative Msg5 expression levels), RIV values are increase in a graded response, but above a threshold of Msg5 cells plateau to high RIV values (above 0.1). High RIV values indicate the difference between inner and outer cell growth as measure by intensity is low. Low RIV values indicate a large difference between inner and outer intensity values (see Methods for calculation of RIV). "W" indicates wild-type. "T" indicates transitory region.

We used the calculated values of Msg5 expression to plot the pathway activation curve as described for the positive network diverter. At low Msg5 levels, a steady

decrease in pathway activity is observed with increasing Msg5 (Figure 2.13). Within a narrow range of expression the response curve is very sensitive to changes in Msg5 expression, supporting the observed requirement for switches with stringent control over basal levels to build negative diverters that achieve conditional rerouting of cell fate. To facilitate comparison between both expression modes, pathway activity values were calculated based on the high GFP population. Inclusion of the low GFP population showed similar results for the transitory range and pathway response curve sensitivity (Supplementary Figure 2.10). While we suspect that the pathway response curve for diverters incorporating constitutive expression is more sensitive to changes in Msg5 than that for feedback diverters, the data do not provide sufficient resolution for differentiating features of the curves that may vary between the two expression modes.

We have demonstrated that negative molecular network diverters can conditionally attenuate pathway activity in the yeast mating pathway, routing cells to the chaste fate. Extending the construction of negative diverters to mammalian cells offers the potential to intervene in the establishment and proliferation of cancers by counteracting aberrant pathway activation. Applying these circuits to diseases such as cancer may reduce or eliminate the proliferation of tumors by checking uncontrolled growth. Additionally, a negative molecular network diverter may be able to restrict pathway attenuation based on temporal and spatial distribution of the environmental trigger limiting the potential side-effects associated with applying such diverters as therapeutics.

Discussion

We have demonstrated the construction of molecular network diverters that positively and negatively regulate pathway activity, conditionally routing cells to alternative fates in response to specific molecular inputs. To construct these diverters, we identified titratable regulators of pathway activity and traced the pathway response curve to regulator levels, identifying the transitory range between the wild-type and alternative fates. In exploring different network architectures, we demonstrated that positive feedback diverters enhance ultrasensitivity in the native pathway, shifting the transitory range for Ste4 to lower levels and providing more robust phenotypic switching. The potency of the negative regulator Msg5 required stringent control of the basal levels for maintenance of the wild-type phenotype in the absence of the environmental trigger. Our studies also demonstrated that these synthetic circuits can be rationally tuned by the exchange of well-characterized genetic parts.

We have demonstrated that within multiple yeast MAPK pathways there are titratable regulators of pathway activity. Ectopic overexpression of specific pathway regulators at these regulatory nodes can be used to modulate pathway activity, providing a readily implementable, modular interface between synthetic regulatory systems and these native pathways. The identification of titratable regulators in multiple pathways suggests that our method for identifying regulators of pathway activity may be broadly applicable to MAPK pathways, facilitating the construction of synthetic circuits that interface with and modulate the function of an important class of signaling pathways. While previous work supports the identified components as regulators of pathway activity [10, 14, 17], we demonstrate that within the mating pathway overexpression of

these regulators above a particular threshold can route to an alternative fate. Additionally, by tracing the pathway response curve in our studies, we have identified the range of regulator expression over which fate transitions from wild-type to the programmed alternative fate. Interestingly, we observed that the transition is reached at lower expression levels for Msg5 than for Ste4, indicating that Msg5 is a more potent regulator of pathway activity. Since Ste4 overexpression leads to cell-cycle arrest, weaker potency of this regulator may represent a control scheme to compensate for variable expression that might otherwise decrease population fitness by reducing overall growth rates. Since Msg5 does not inhibit growth, limiting its potency may be less critical to population fitness.

The pathway response curve is unique to its regulator and dependent on the molecular details of the regulator's interaction with other components in the native network. While modifying the regulator's interaction via protein engineering can change these interactions to reshape the pathway response [8, 17], changing the connectivity of a synthetic circuit can provide a simpler method for rationally reshaping the pathway response curve. In our system, we demonstrated that the introduction of positive feedback shifts the threshold at which cell fate diverges to lower relative Ste4 expression levels, improving phenotypic switching and suggesting an enhancement of the pathway's ultrasensitivity. Further, positive autoregulatory loops such as these have been shown to exhibit bistability [8, 21]. When properly tuned, positive feedback loops can thus provide a providing a type of memory to fate commitment by sustaining the network response when triggered above a particular input level [23]. While the introduction of positive feedback amplifies phenotypic switching from the network diverter, negative feedback

plays a role in reducing population heterogeneity. In our system we observed that the introduction of negative feedback into the network diverter reduced population heterogeneity, resulting in a single narrow peak of pathway activity values. Feedback expression provides a negative response that is proportional to pathway activity which may asymptote to a particular stable state. Mismatches in pathway activity and Msg5 expression distributions are attenuated for feedback expression. In contrast, constitutive expression results in a more broadly distributed response in which mismatches in pathway expression and regulator expression levels introduced by extrinsic noise increase population heterogeneity. Thus, the introduction of feedback architectures into the molecular network diverters result in synthetic circuits that exhibit useful properties for robust circuit performance, including enhanced ultrasensitivity and reduced population heterogeneity.

We also demonstrate that molecular network diverters can be rationally tuned by the exchange of well-defined genetic parts. Specifically, the promoter and switch elements can be exchanged within the diverter architecture to predictably tune pathway activity. In addition, we showed that the ability of the switch component within the diverter to route cell fate is independent of the input signal the switch is programmed to detect and dependent on the quantitative gene-regulatory activities exhibited by the switch. This property allows for orthogonal sets of network diverters, which can be triggered by independent input channels, to be designed through the incorporation of RNA switches responsive to different molecular inputs. Further, independent channels of regulation provide for the integration of multiple network diverters that control different pathway nodes, allowing independent triggering of each diverter. This capability can

ultimately be used to route genetically identical cells to different cell fates dependent on the small-molecule trigger present in the environment. Such orthogonal control strategies may facilitate patterning of multiple cell fates from a genetically homogenous culture via the controlled distribution of small-molecule triggers.

Our studies also highlight the design trade-offs and considerations in configuring network diverters to span the range of expression required for fate routing. While RNA switches with higher SAR values exhibit greater fold changes in regulator expression, regulators that exhibit high activities, such as Msg5, require more stringent switches with lower basal levels that can exhibit smaller dynamic ranges (SAR). In certain situations, the optimal diverter configuration and fate switching may not be achievable based on the pathway response curve and the set of available parts for composing the diverter. Therefore, the development of strategies to amplify input-triggered switching and an expanded the set of switches (and corresponding range of properties) will extend the number of native pathways to which molecular network diverters can be applied. In addition, the switch component provides a promoter-independent strategy for tuning gene expression, potentially providing greater flexibility to the control system, particularly in situations in which the application constrains the choice of promoter. For example, synthetic systems developed as therapies for the treatment of diabetes and gout require specific pathway-responsive promoters [24, 25]. In addition to facilitating tuning in these synthetic systems, the RNA switches will enable the construction of synthetic networks that provide exogenous spatial and temporal regulation, providing a strategy for limiting the extent of side-effects associated with such therapies.

The molecular network diverter architecture developed in this study offers the potential to independently trigger a positive and negative diverter via independent small-molecule input channels, allowing routing to three different fates from a genetically identical cell population. However, the integration of both positive and negative diverters into a single cell poses challenges in maintaining the performance of each diverter. While each diverter runs quiescently with respect to fate routing in the absence of small-molecule input, addition of the diverter measurably changes the basal activity of the pathway. Since the opposing diverters perform opposite functions on pathway activity that are processed through the native network, basal expression levels from the non-triggered diverter may antagonize routing from the triggered diverter. The incorporation of RNA switches exhibiting low basal levels and large dynamic ranges may facilitate improved routing in such system designs. However, computational models of RNA switches have indicated a tradeoff in tuning the stringency of a switch and its sensitivity to the input ligand, ultimately impacting the dynamic range [26]. Network architectures that amplify small-molecule triggered routing may enhance fate switching even in the presence of diverter antagonization. Issues with basal expression levels impinging on circuit performance have observed in synthetic network construction [23, 27, 28]. Researchers have addressed such issues by adding layers of post-transcriptional control, such as through riboregulators in bacteria [29] and microRNAs in mammalian cells [30, 31], to reduce gene expression leakage in the OFF state. Thus, RNA control elements that target the opposite regulator may be coexpressed with pathway regulators or imbedded within the opposite transcript to achieve reduced basal level leakage, diminished diverter

antagonization, and amplified diverter-mediated switching when positive and negative diverters are simultaneous implemented.

While our work elucidates the design principles for constructing molecular network diverter systems in yeast, these principles can be extended to construct similar systems in higher eukaryotes, offering the potential to probe native networks in multicellular organisms, program tissue formation, and target cancer therapies with synthetic molecular network diverters. Additionally, these synthetic circuits may provide the control tools necessary for the *ex vivo* construction of complex tissues when combined with designer scaffolds printed with small-molecule triggers. Further, our work highlights the utility of control systems that can be conditionally induced while maintaining endogenous connectivity with independent promoter selection. For genetic programs that decrease fitness or viability, such systems allow for conditional induction at the appropriate time, reducing the selection pressure to eliminate the prescribed program. As RNA switches are generated to sense a broader array of molecular ligands, diverters may be configured to respond to endogenous inputs, potentially allowing autonomous control within the cell. As our ability to build interfaces between synthetic circuits and native networks improves, synthetic biological devices that control cell fate will translate into therapeutic and regenerative medicine applications, as well as facilitate the study of natural biological systems.

Materials and methods

Plasmid and strain construction

Standard molecular biology cloning techniques were used to construct all plasmids [32]. DNA synthesis was performed by Integrated DNA Technologies (Coralville, IA). All enzymes, including restriction enzymes and ligases, were obtained through New England Biolabs (Ipswich, MA). Ligation products were electroporated with a GenePulser XCell (Bio-Rad, Hercules, CA) into an *E. coli* DH10B strain (Invitrogen, Carlsbad, CA), where cells harboring cloned plasmids were maintained in Luria-Bertani media containing 50 mg/ml ampicillin (EMD Chemicals). All cloned constructs were sequence verified by Laragen Inc (Santa Monica, CA).

Plasmids expressing GFP were constructed from pCS321 (Supplementary Figure 2.11A, Supplementary Table 2.1 [33]. Promoters pLOW (pADH1), pMED (pCYC1), pHIGH (TEF1 mutant 7) [34], pHIGHEST (pTEF1), pSTL1, and pFUS1 were PCR amplified and cloned into SacI (or NotI) and BamHI in pCS321 using promoter-specific primers (Supplementary Table 2.2).

Plasmids for galactose-titration were constructed from pCS1128 a TRP plasmid which bears pGAL1-yEGFP-CYCt with an ON state control cloned in the 3'UTR between AvrII and XhoI (Supplementary Figure 2.11B, Supplementary Table 2.3). Mating genes (Ste4, Ste50, Ste11, Ste7, Fus3, Msg5) and osmo genes (Pbs2, Hog1, Ptc1) were cloned between AvrII and BamHI using gene-specific primers for amplification from plasmid and genomic templates (Supplementary Table 2.2). XhoI restriction sites in Fus3 and Ste4 and an AvrII restriction site in Ste7 were removed by site-directed mutagenesis using QuikChange II Site-directed mutagenesis kit (Agilent, Santa Clara,

CA) according to the manufacturer's instruction with appropriate primers (Supplementary Table 2.4). Pbs2 was cloned into pCS321 using the same primers and restriction sites to generate a URA plasmid Pbs2 titration construct.

To construct the negative and positive molecular network diverters, Msg5 and Ste4 were cloned into the GFP expressing constructs with the various promoters as described above for the galactose titration constructs. Ribozyme-based devices and appropriate controls (Supplementary Table 2.5) were inserted into the 3' UTR via the unique restriction sites AvrII and XhoI, located immediately downstream of the gene stop codon as described previously (Supplementary Tables 3.6 and 3.7) [33].

Reporter strains for galactose-titration studies were constructed by knocking out the galactose transporter, GAL2, to construct the mating strain CSY532 and osmolarity strain CSY139. The pFUS1-yEGFP3-CYC1t cassette from pCS1124 was cloned into pCS1391, a loxP integrating vector [35], via the unique restriction sites SacI and KpnI to make pCS2292 (Supplementary Figure 2.12, Supplementary Table 2.8). pFUS1-yEGFP3-CYC1t was chromosomally integrated into yeast strain CSY364 (EY1119; W303a Δsst1 Δkss1::HIS3) [36] via homologous recombination using the gene cassette from pCS2292 to construct yeast strain CSY532 (W303aΔsst1 Δkss1::HIS3 gal2 ::FUS1p-yEGFP3-loxP-KanR) (Supplementary Table 2.9). Briefly, the pFUS1-yEGFP3-loxP-KanR cassette was PCR amplified using Expand High Fidelity PCR system (Roche, Indianapolis, IN) from pCS2292 using forward and reverse primers (Supplementary Table 2.10), each carrying 60 nts of homologous sequence to the *GAL2* locus. Yeast strain EY1119 was transformed with 12 μg of gel purified PCR product and plated on G418 plates to build yeast strain CSY532. CSY408 was constructed as described above using a blank

integration cassette. To construct a reporter for the osmolarity pathway, pSTL1-yEGFP-CYC1t was cloned into a yeast disintegrator plasmid pCS1441 which is specific to the FCY1 locus via unique restriction site KpnI and SacI. CSY408 was was transformed with disintegrator vector bearing pSTL1-yEGFP-CYC1t linearized by digestion with AscI as described previously [37] to yield KGY139. For characterization of the molecular network diverters, yeast strain CSY840 was constructed as described for CSY532 using primers specific to the *TRP1* locus (Supplementary Tables 3.9 and 3.10).

Measuring mating pathway activity via a transcriptional reporter

For the galactose titration studies, plasmids bearing the galactose-inducible promoter (pGAL1) controlling expression of various pathway regulators and appropriate controls were transformed into yeast strain CSY532. Cells were inoculated into the appropriate dropout media, grown overnight at 30°C, and back diluted into fresh media in the presence of varying concentrations of galactose (0, 0.25, 0.5, 1, 2 and 3%) to an OD_{600} of < 0.1. To identify negative pathway regulators, after growing for 3 hr at 30°C, cells were stimulated with saturating pheromone levels, to a final concentration of 100 nM α mating factor acetate salt (Sigma-Aldrich, St. Louis, MO), to activate the mating pathway. The pathway was not stimulated when evaluating the potential of positive regulators. Following 6 hr of growth post-back-dilution, GFP fluorescence levels from the pFUS1-yEGFP3 reporter were evaluated via flow cytometry using a Cell Lab Quanta SC flow cytometer (Beckman Coulter, Fullerton, CA) with the following settings: 488-nm laser line, 525-nm bandpass filter, and photomultiplier tube setting of 5.0 on FL1 (GFP). Fluorescence data were collected under low flow rates for ~ 10,000 viable cells.

Normalized pathway activity is calculated as the geometric mean of three biological replicates of each sample normalized to the blank plasmid control not bearing a gene at the corresponding galactose concentration.

The molecular diverter plasmids and appropriate controls were transformed into yeast strain CSY840. Cells were inoculated into the appropriate drop-out media, grown overnight at 30°C, and back diluted into fresh media in the presence or absence of ligand at the specified concentration to an OD_{600} of < 0.1. For negative diverters, after growing for 3 hr at 30°C, cells were stimulated with saturating pheromone levels, to a final concentration of 100 nM α mating factor acetate salt, to activate the mating pathway. Following 6 hr of growth post-back-dilution, GFP fluorescence levels from the pFUS1-yEGFP3 reporter were evaluated via flow cytometry using a Cell Lab Quanta SC flow cytometer as described previously. Normalized pathway activity is calculated as the geometric mean of three biological replicates of each sample normalized to the blank plasmid control stimulated with saturating α mating factor in the absence of either small molecule.

Measuring osmolarity pathway activity via a transcriptional reporter

Galactose titration studies in the osmolarity pathway were performed as described previously for the mating pathway with the following exceptions. Cells were transformed in KGY139 to meaure the activity of the osmolarity pathway via the integrated pSTL1-GFP reporter. Normalized pathway activity is calculated relative to the pathway stimulated with 1M sorbitol for 6 hr. Quanta flow cytometer photomultiplier tube setting of 7.5 on FL1 (GFP).

Measuring mating pathway activity via halo assays

Mating associated cell-cycle arrest was evaluated via halo assays [38]. Halo assays were performed on cultures grown overnight in YNB with appropriate dropout solution, back diluted into fresh media, and grown to OD_{600} ~ 0.2–0.4. 200 μl of each replicate was plated on the appropriate drop-out plates. For galactose-titration halo assays, cells were plated on noninducing, nonrepressing plates with varying concentrations of galactose (0, 0.25, 0.5, 1, 2 and 3%). For characterization of the molecular network diverters, cells were plated on plates containing no small molecule or 5 mM theophylline or 1 mM Tetracycline as specified. After plating the cells, a gradient of α mating factor was established by saturating a filter disk (2 mm diameter) of Whatman paper with 9 μl of 0.1 mg/mL α mating factor and placing the disk on the center of the plate. Cells were grown for 18 hr at 30°C and imaged via epi-white illumination with a GelDoc XR+ System (Bio-Rad).

Relative intensity values (RIV) were calculated to measure the difference in inner and outer halo cell growth. Intensity of rectangular cross-sections at r = 0.15 inches (inner region) and r = 0.5 inches (outer region) from the filter disc were calculated using image processing software from the GelDoc XR+ System. The difference between outer and inner intensity values were calculated for each plate to yield the plate's intensity range. The difference between each plate intensity range and max intensity range was calculated and normalized by the max intensity range. The max intensity range was calculated by subtracting the minimal intensity value from the maximal intensity value for the entire set

of plates including both nonswitch controls and the blank plasmid control. Increasing RIV values indicate decreasing differences in inner and outer intensities.

Measuring mating pathway activity via Western blots

Yeast cells harboring appropriate constructs were grown as indicated above in 5 mL cultures to ~ 1.0 OD_{600}, pelleted via centrifugation, and resuspended in 800 ul modified RIPA Lysis buffer (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1% Triton X-100, 0.5% Nonidet P-40, 0.25% sodium deoxycholate, 50 mM β-glycerophosphate (pH 7.3), 10 mM NaPP, 30 mM NaF, 1 mM benzamidine, 2 mM EGTA, 1 mM sodium orthovanadate, 1 mM dithiothreitol, 5 μg/ml aprotinin, 5 μg/ml leupeptin, 1 μg/ml pepstatin and 1 mM phenylmethylsulfonyl fluoride). The resuspended cell mixture was added to a tube containing 1 mm glass beads and shaken for 5 min at 30/sec in a homogenizer to generate cell lysates. Following homogenization, the lysates were boiled for 5 minutes in a 95 ^{o}C sand bath and subsequently centrifuged at maximum speed for 2 min. The supernatant was recovered into a fresh tube. Cleared lysates were stored at -80^{o}C.

Sample analysis was performed by adding one part 10X SDS sample buffer (50 mM Tris–HCl (pH=6.8), 12% (vol/vol) glycerol, 2% (wt g/vol ml) SDS, 1% DTT, 0.01% (wt g/vol ml)bromophenol blue) to 9 parts of the cell lysate. Samples were heated at 100^{o}C for 5 min and subsequently loaded on a 10% acrylamide SDS-polyacrylamide gel. 3 μL of SeeBlue Plus2 Prestained Standard (Invitrogen) was loaded into a separate lane as a size marker. Different volumes of lysate and loading buffer were added to each lane up to a total volume of 30 μL to adjust for equal protein loading. The gel was run at 50

mA for ~ 1 hr and transferred in transfer buffer and apparatus to nitrocellulose. Immunoblotting and imaging was performed as previously described using an anti-phospho P44/42 antibody (Cell Signaling Technology to detect phosphorylated Fus3 and anti-Cdc28 as a loading control (Sc-53, Santa Cruz Biotechnology, Inc, Santa Cruz, CA) [39].

Calculating relative Ste4 and Msg5 expression levels

For constitutive expression of Ste4 and Msg5, each construct's promoter strength (Supplementary Figure 2.2) was multiplied by the switch percent expression at the appropriate ligand concentration. Each value was normalized to expression calculated from pHIGH-ON (the highest expression in these assays). For feedback expression, normalized pathway activity values (pFUS1-GFP) were taken as a proxy for the strength of pFUS1-Ste4 or pFUS1-Msg5 expression. These values were evaluated as a measure the promoter strength and, as for constitutive expression, multiplied by the switch percent expression at the appropriate ligand concentration. Each value was normalized to expression calculated from pHIGH-ON.

Acknowledgements

Special thanks to Richard Yu for technical advice for halo assays, Stephen Chapman for technical assistance with mating assays and Western blots, and the Van Drogen lab for constructs. Funding provided by the National Institutes of Health (R01GM086663).

References

1. Seger R, Krebs EG: **The MAPK signaling cascade.** *Faseb J* 1995, **9:**726-735.
2. Hanahan D, Weinberg RA: **The hallmarks of cancer.** *Cell* 2000, **100:**57-70.
3. McCormick F: **Signalling networks that cause cancer.** *Trends Cell Biol* 1999, **9:**M53-56.
4. Discher DE, Mooney DJ, Zandstra PW: **Growth factors, matrices, and forces combine and control stem cells.** *Science* 2009, **324:**1673-1677.
5. Lutolf MP, Gilbert PM, Blau HM: **Designing materials to direct stem-cell fate.** *Nature* 2009, **462:**433-441.
6. Chapman S, Asthagiri AR: **Resistance to signal activation governs design features of the MAP kinase signaling module.** *Biotechnol Bioeng* 2004, **85:**311-322.
7. Levchenko A, Bruck J, Sternberg PW: **Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties.** *Proc Natl Acad Sci U S A* 2000, **97:**5818-5823.
8. Wang X, Hao N, Dohlman HG, Elston TC: **Bistability, stochasticity, and oscillations in the mitogen-activated protein kinase cascade.** *Biophys J* 2006, **90:**1961-1978.
9. Andersson J, Simpson DM, Qi M, Wang Y, Elion EA: **Differential input by Ste5 scaffold and Msg5 phosphatase route a MAPK cascade to multiple outcomes.** *Embo J* 2004, **23:**2564-2576.
10. Ingolia NT, Murray AW: **Positive-feedback loops as a flexible biological module.** *Curr Biol* 2007, **17:**668-677.

11. Santos SD, Verveer PJ, Bastiaens PI: **Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate.** *Nat Cell Biol* 2007, **9:**324-330.

12. Win MN, Smolke CD: **A modular and extensible RNA-based gene-regulatory platform for engineering cellular function.** *Proc Natl Acad Sci U S A* 2007.

13. Palpant RG, Steimnitz R, Bornemann TH, Hawkins K: **The Carter Center Mental Health Program: addressing the public health crisis in the field of mental health through policy change and stigma reduction.** *Prev Chronic Dis* 2006, **3:**A62.

14. Bashor CJ, Helman NC, Yan S, Lim WA: **Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics.** *Science* 2008, **319:**1539-1543.

15. Varghese J, Cohen SM: **microRNA miR-14 acts to modulate a positive autoregulatory loop controlling steroid hormone signaling in Drosophila.** *Genes Dev* 2007, **21:**2277-2282.

16. Herranz H, Cohen SM: **MicroRNAs and gene regulatory networks: managing the impact of noise in biological systems.** *Genes Dev* 2010, **24:**1339-1344.

17. Park SH, Zarrinpar A, Lim WA: **Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms.** *Science* 2003, **299:**1061-1064.

18. Mettetal JT, Muzzey D, Gomez-Uribe C, van Oudenaarden A: **The frequency dependence of osmo-adaptation in Saccharomyces cerevisiae.** *Science* 2008, **319:**482-484.

19. Nevozhay D, Adams RM, Murphy KF, Josic K, Balazsi G: **Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression.** *Proc Natl Acad Sci U S A* 2009, **106:**5123-5128.

20. Xiong W, Ferrell JE, Jr.: **A positive-feedback-based bistable 'memory module' that governs a cell fate decision.** *Nature* 2003, **426:**460-465.

21. Ferrell JE, Jr.: **Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability.** *Curr Opin Cell Biol* 2002, **14:**140-148.

22. Becskei A, Serrano L: **Engineering stability in gene networks by autoregulation.** *Nature* 2000, **405:**590-593.

23. Ajo-Franklin CM, Drubin DA, Eskin JA, Gee EP, Landgraf D, Phillips I, Silver PA: **Rational design of memory in eukaryotic cells.** *Genes Dev* 2007, **21:**2271-2276.

24. Han J, McLane B, Kim EH, Yoon JW, Jun HS: **Remission of diabetes by insulin gene therapy using a hepatocyte-specific and glucose-responsive synthetic promoter.** *Mol Ther* 2011, **19:**470-478.

25. Kemmer C, Gitzinger M, Daoud-El Baba M, Djonov V, Stelling J, Fussenegger M: **Self-sufficient control of urate homeostasis in mice by a synthetic circuit.** *Nat Biotechnol* 2010, **28:**355-360.

26. Beisel CL, Smolke CD: **Design principles for riboswitch function.** *PLoS Comput Biol* 2009, **5:**e1000363.

27. Gardner TS, Cantor CR, Collins JJ: **Construction of a genetic toggle switch in Escherichia coli.** *Nature* 2000, **403:**339-342.

28. Friedland AE, Lu TK, Wang X, Shi D, Church G, Collins JJ: **Synthetic gene networks that count.** *Science* 2009, **324:**1199-1202.

29. Callura JM, Dwyer DJ, Isaacs FJ, Cantor CR, Collins JJ: **Tracking, tuning, and terminating microbial physiology using synthetic riboregulators.** *Proc Natl Acad Sci U S A* 2010, **107:**15898-15903.

30. Deans TL, Cantor CR, Collins JJ: **A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells.** *Cell* 2007, **130:**363-372.

31. Xie Z, Wroblewska L, Prochazka L, Weiss R, Benenson Y: **Multi-input RNAi-based logic circuit for identification of specific cancer cells.** *Science* 2011, **333:**1307-1311.

32. Sambrook J RD: *Molecular Cloning: A Laboratory Manual, 3rd edn.* Cold Spring Harbor, NY: Cold Spring Harbor Lab Press; 2001.

33. Win MN, Smolke CD: **A modular and extensible RNA-based gene-regulatory platform for engineering cellular function.** *Proceedings of the National Academy of Sciences* 2007, **104:**14283-14288.

34. Nevoigt E, Kohnke J, Fischer CR, Alper H, Stahl U, Stephanopoulos G: **Engineering of promoter replacement cassettes for fine-tuning of gene expression in Saccharomyces cerevisiae.** *Appl Environ Microbiol* 2006, **72:**5266-5273.

35. Guldener U, Heck S, Fielder T, Beinhauer J, Hegemann JH: **A new efficient gene disruption cassette for repeated use in budding yeast.** *Nucleic Acids Res* 1996, **24:**2519-2524.

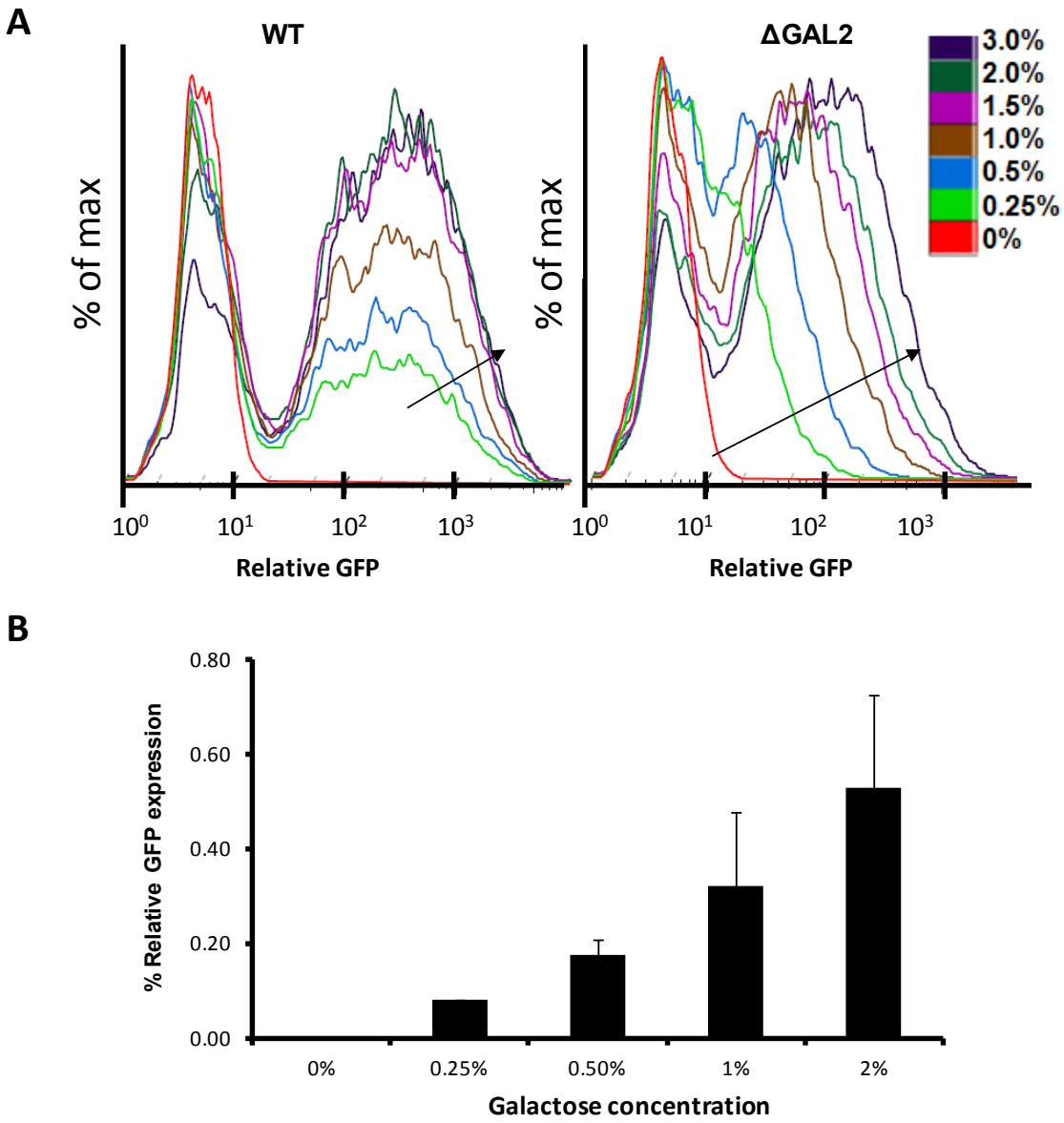
36. Flotho A, Simpson DM, Qi M, Elion EA: **Localized feedback phosphorylation of Ste5p scaffold by associated MAPK cascade.** *J Biol Chem* 2004, **279:**47391-47401.

37. Sadowski I, Su TC, Parent J: **Disintegrator vectors for single-copy yeast chromosomal integration.** *Yeast* 2007, **24:**447-455.

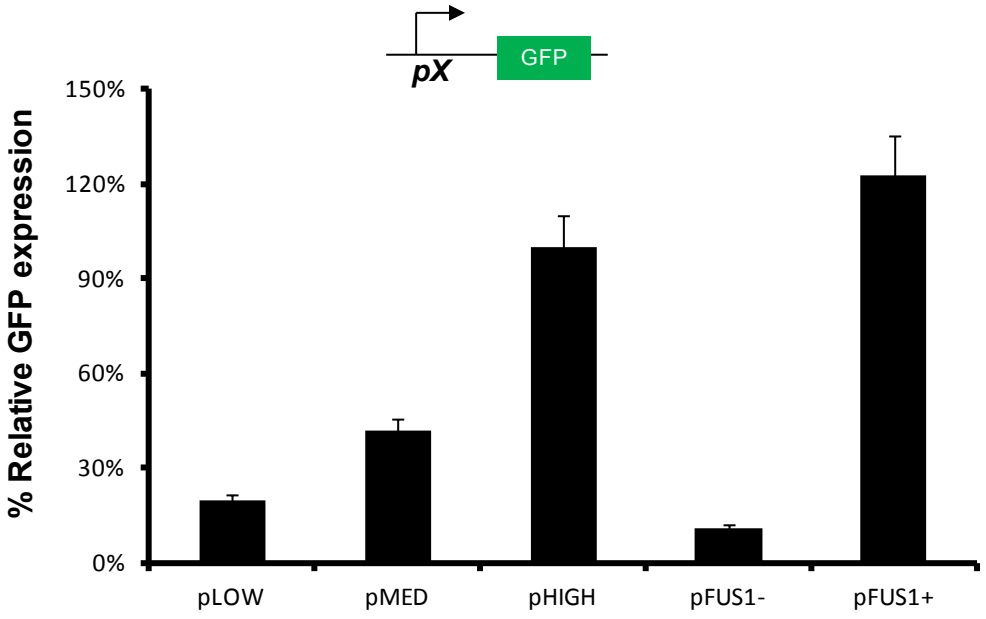
38. Sprague GF, Jr.: **Assay of yeast mating reaction.** *Methods Enzymol* 1991, **194:**77-93.

39. Chapman SA, Asthagiri AR: **Quantitative effect of scaffold abundance on signal propagation.** *Mol Syst Biol* 2009, **5:**313.

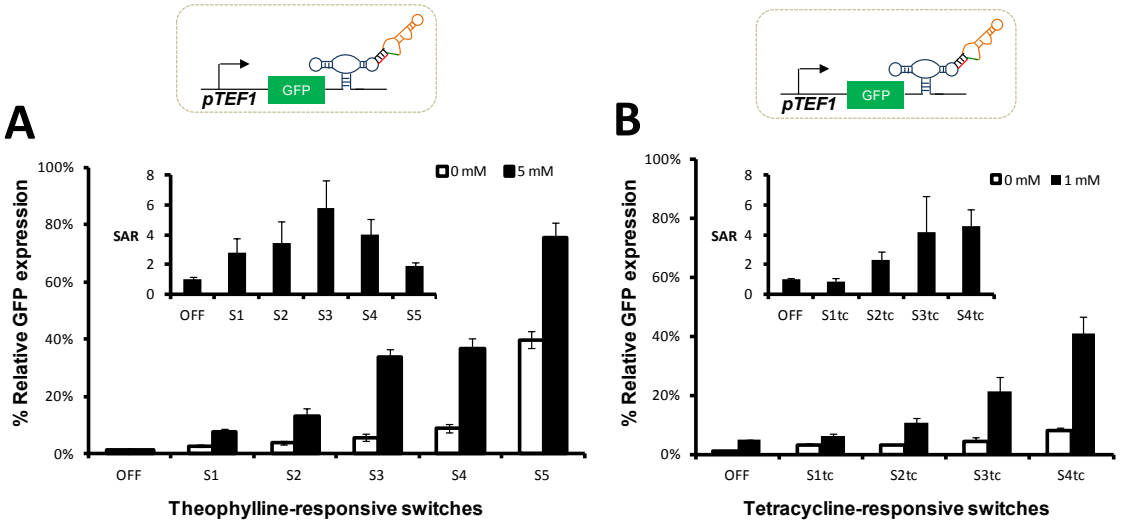
Supplementary figures



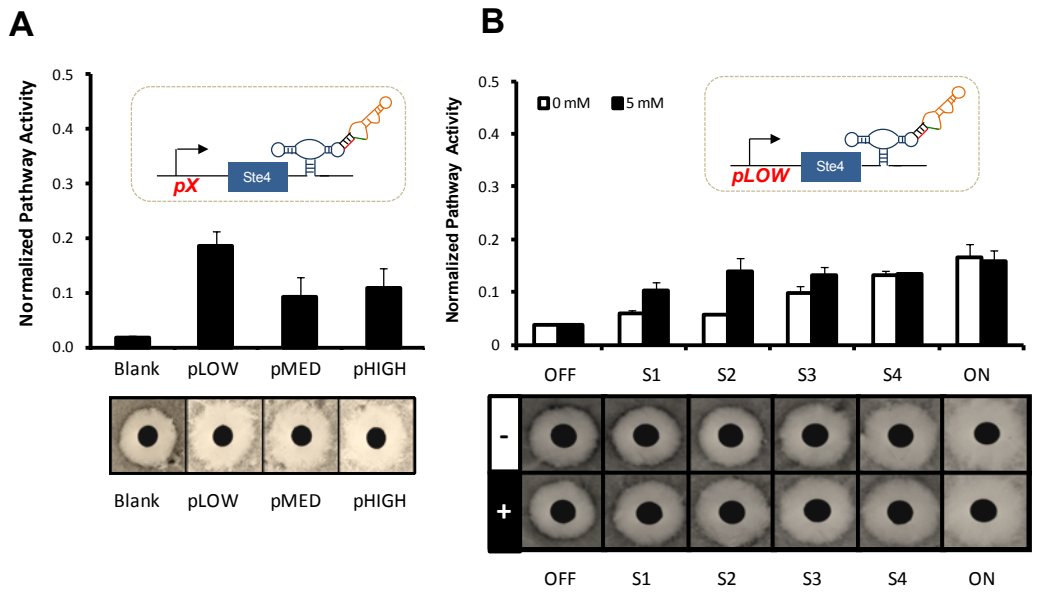
Supplementary Figure 2.1. Engineered strain linearly increases the mean level of expression from the galactose-inducible promoter in response to increasing galactose. A. Histograms for the wild-type (WT) strain show that the mean levels of expression do not significantly change in response to increasing levels of galactose. In the engineered strain (ΔGAL2), histograms demonstrate that the mean fluorescence increases linearly with galactose concentration. In both strains, the percentage of cells in the high expression population increases as galactose levels increase. Arrows indicate direction of increasing galactose concentration. Cellular fluorescence was measured via flow cytometry 6 hours post-induction with galactose from cells harboring a construct bearing pGAL1-Fus3-GFP. Histograms are from single samples and are representative of three replicates. **B.** Mean fluorescence levels of GFP from cells bearing a plasmid with pGAL1-GFP. Assays were performed as described previously. Data represents mean of three replicates with the 0% (noninduced) control subtracted and values normalized to pHIGH-GFP expression levels. Standard deviation is used for the error bars.



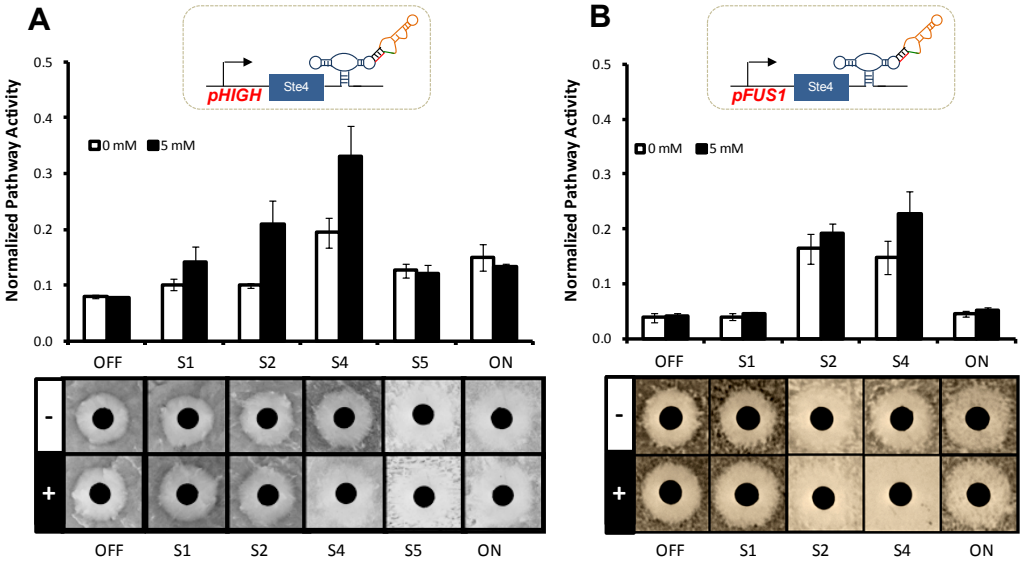
Supplementary Figure 2.2. Promoter characterization. Promoter characterization is critical for placing the regulators at the appropriate levels of expression. Mean values of viable cells were calculated and normalized by the value of pHIGH. Not shown on the graph is pHIGHEST which is more than 6 times higher than pHIGH (reference TEF1 library). pLOW refers to pADH1, pMED to pCYC1, pHIGH to pTEF1mutant7 and pHIGHEST to pTEF1. pFUS1 is the mating-responsive promoter. GFP values were determined in the absence of α factor (pFUS1-) and at saturating concentration (100 nM , pFUS1+) three hours following stimulation.



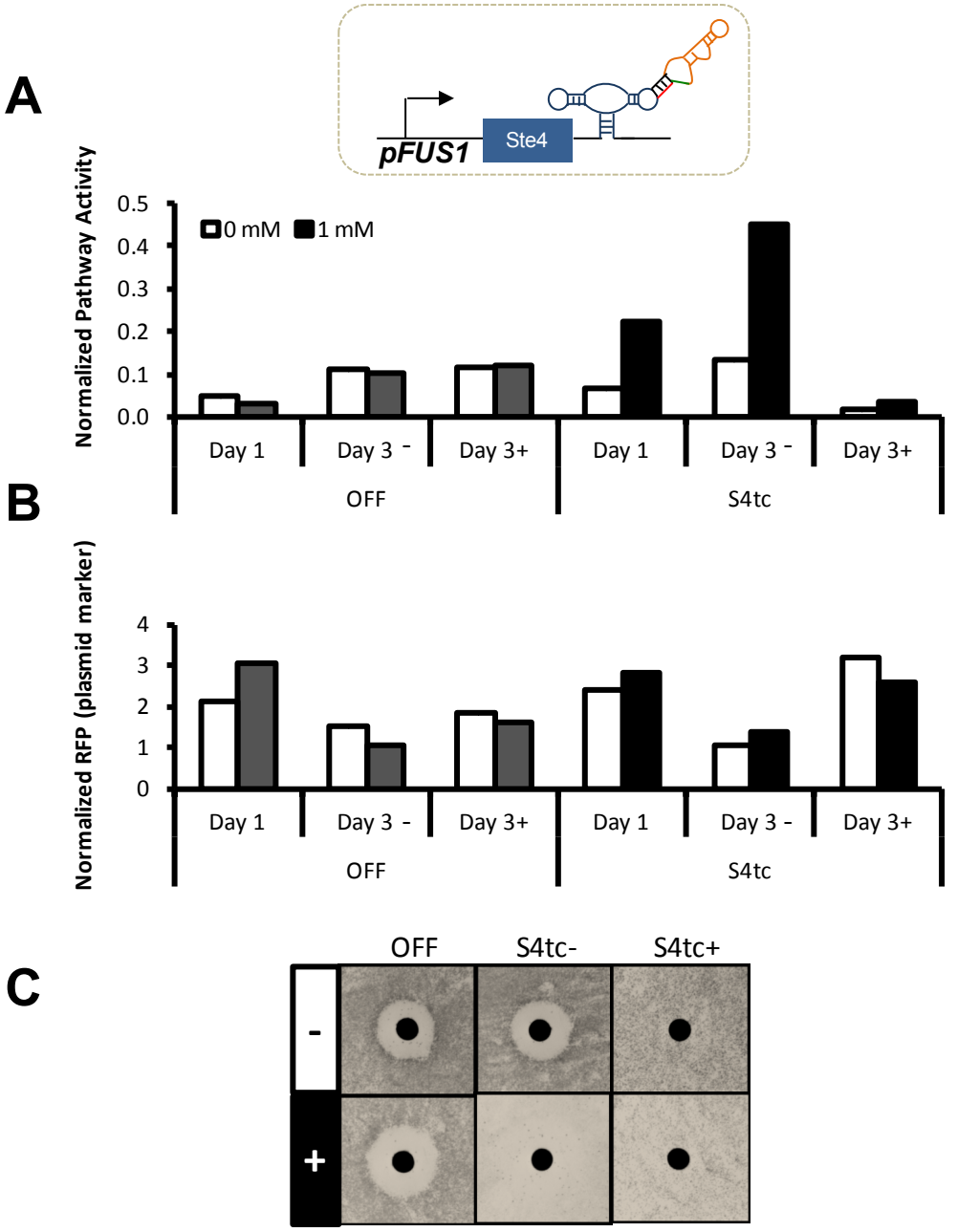
Supplementary Figure 2.3. Range of switch expression strengths. A. Theophylline-responsive switches' basal levels, the expression levels in the absence of ligand, range from 3% to 40% of the ON control. S3 has the highest switch activation ratio (SAR) at 5.7 . SAR is the ratio of expression levels in the presence of ligand to the level in the absence of ligand. **B.** Tetracycline-responsive switches' basal levels range from 3% to 8%. S3tc and S4tc have similar high SARs at 5.2 and 5.1, respectively.



Supplementary Figure 2.4. Constitutive expression with low-strength promoter shows high pFUS1-GFP levels, yet switches fail to cross the phenotypic transitory range. A. Counterintuitively, higher pFUS1-GFP resulted from Ste4 expression from the LOW promoter compared to either the MED or HIGH promoter. Halo assays demonstrated similar phenotypic outcome across the range. **B.** Ste4 expression from the LOW promoter and paired with a range of theophylline-responsive switches shows weak routing to the promiscuous fate. As switch strength increases, cell growth slowly diminishes but ligand-induced routing to the promiscuous fate is not observed. The threshold to the promiscuous fate is postulated to be above the pLOW-Ste4-ON expression levels.

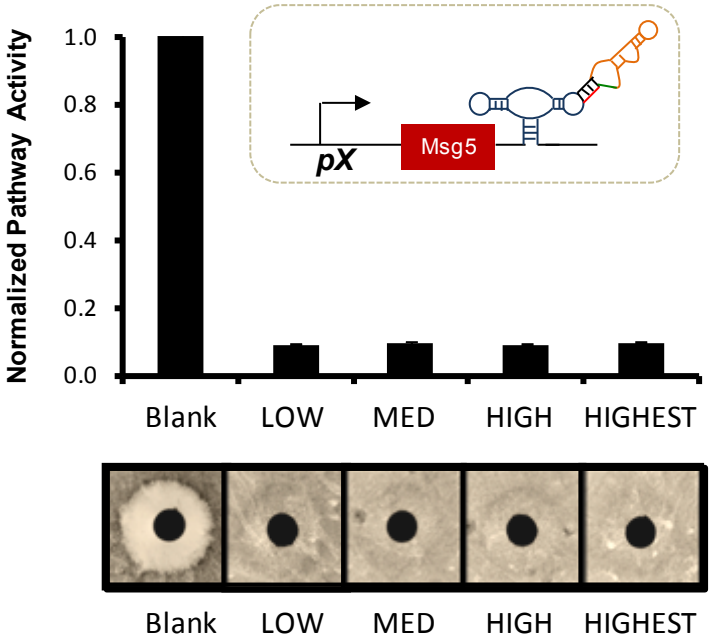


Supplementary Figure 2.5. Theophylline-inducible positive network diverters mirror response of tetracycline-responsive diverters. A. Ste4 expression from pHIGH paired with theophylline-responsive switches shows that pathway activity increases across the range up to S4. Above S3 expression levels, pathway activity as measured by pFUS1-GFP drops. Addition of theophylline increases pathway activity for each switch construct up to S3. For S3, plate assays in the absence of theophylline show a wild-type halo, while at 5 mM theophylline cells adopt the promiscuous phenotype. **B.** Constructing a positive feedback loop amplifies ligand-induce phenotype switching. S3 shows theophylline-induced phenotype switching.

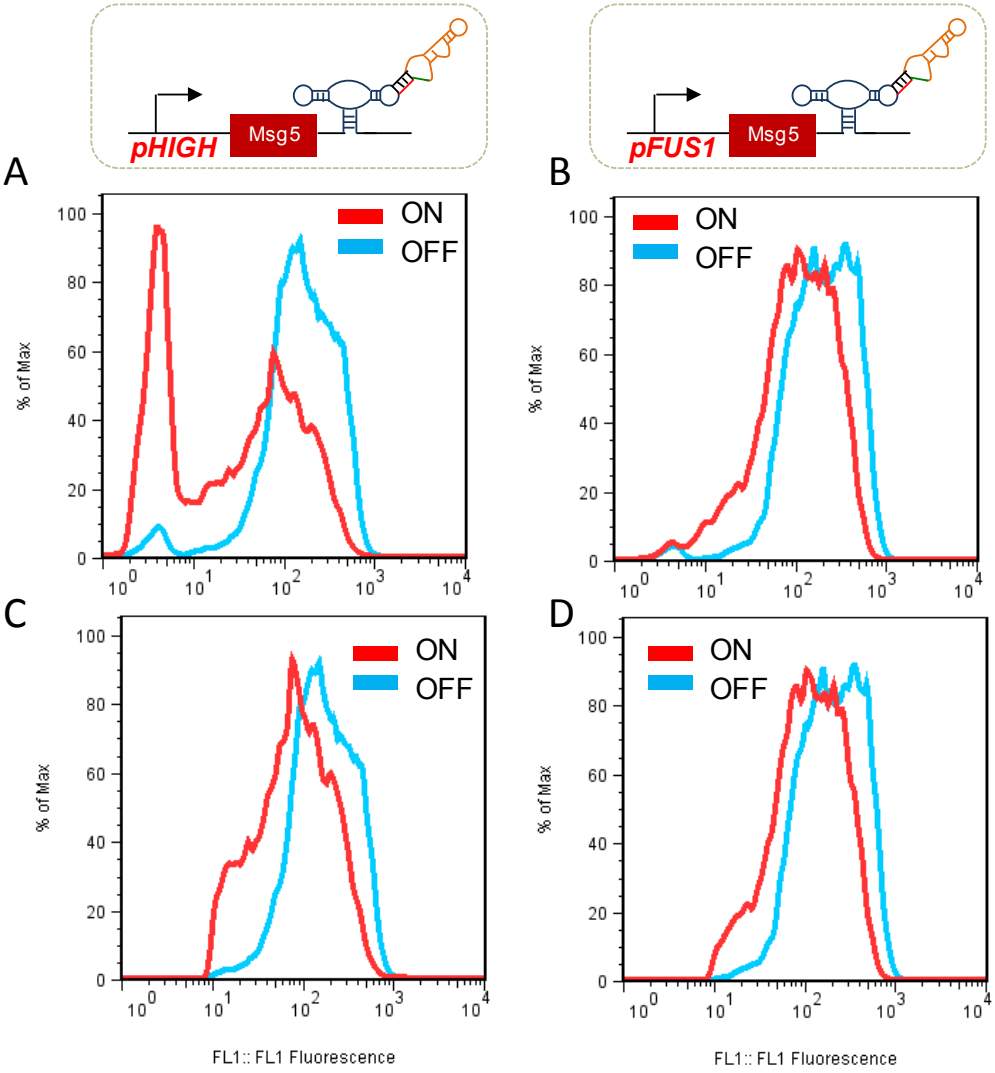


Supplementary Figure 2.6. Selection of mating-resistant cells occurs over time at superthreshold levels of positive feedback. A. Pathway activity drops over time for cells grown at super-threshold levels of positive feedback but not at low levels of positive feedback. Cells expressing the OFF state control with pFUS1-Ste4 show constant, low levels of pathway activity whether grown in the presence (+) or absence (-) of tetracycline for 2 days. S4tc shows normal increases in pathway activity when triggered with tetracycline on Day 1 and Day 3. However, cells triggered after 2 days growth in tetracycline, at superthreshold levels of positive feedback, do not significantly upregulate pathway activity. Instead, these cells adopt a low pFUS1-GFP profile. Cells were inoculated overnight and on Day 1 back-diluted into the indicated concentration of tetracycline. After 6 hours, fluorescence values for GFP were measured by flow cytometry as described previously. Cells were grown in the presence (+) or absence (-) of tetracycline another 18 hours before back-dilution on Day 2. Cells were then grown 24 hours, back diluted in the presence or absence of tetracycline. After 6 hours, fluorescence levels were measured. **B.** Constitutively expressed

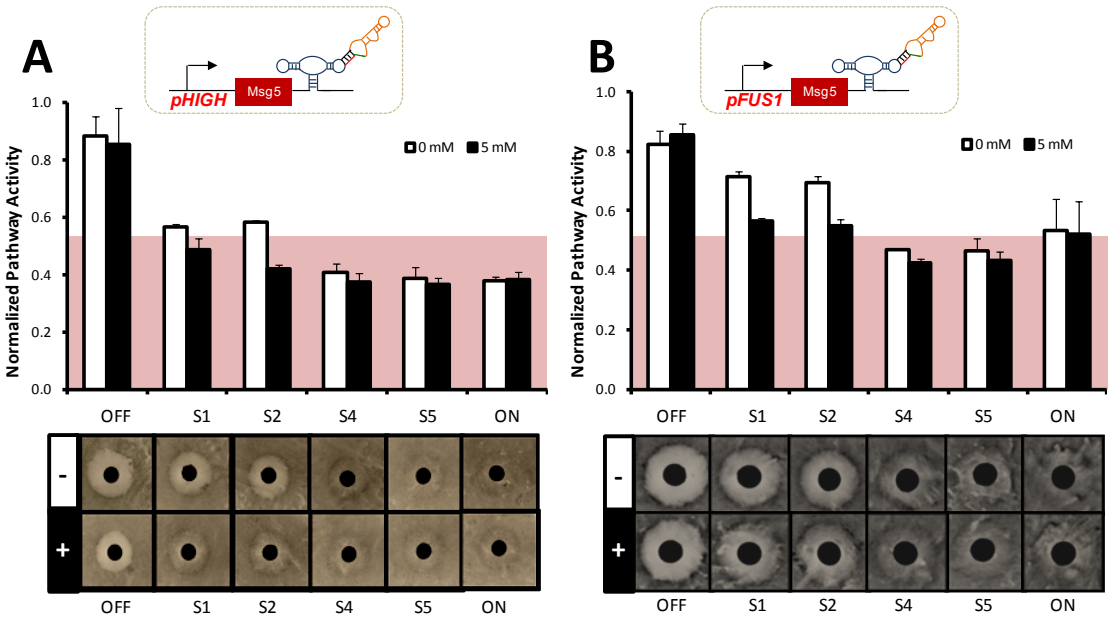
RFP acts as a plasmid marker and indicator of expression from the plasmid. RFP plasmid marker data do not show a correlation between selective pressure and RFP values. Cells experiencing strong selective pressure do not reduce or eliminate expression from the plasmid more than at low selective pressure. These data suggest that mating-resistant cells evade the arrest-inducing program by a method other than jettisoning the plasmid. RFP fluorescence was determined by spectrophotometer plate reader with excitation at 587 nm and emission 610 nm. RFP values were normalized by cells density as measured by absorbance at 600 nm. Final values were computed by dividing by the non-RFP fluorescent control at the corresponding tetracycline concentration. **C.** Halo assays with cells plated in the presence or absence of 1 mM tetracycline after 3 days of growth. pFUS1-Ste4-S4tc shows that cells maintain routing ability after 3 days when grown in the absence of tetracycline (S4tc-). Cells bearing pFUS1-Ste4-S4tc grown in the presence of tetracycline (S4tc+) show resistance to pheromone-induced cell cycle arrest, but still respond to tetracycline by modestly reducing cell growth. The OFF state control maintains normal halo formation in the presence or absence of tetracycline.



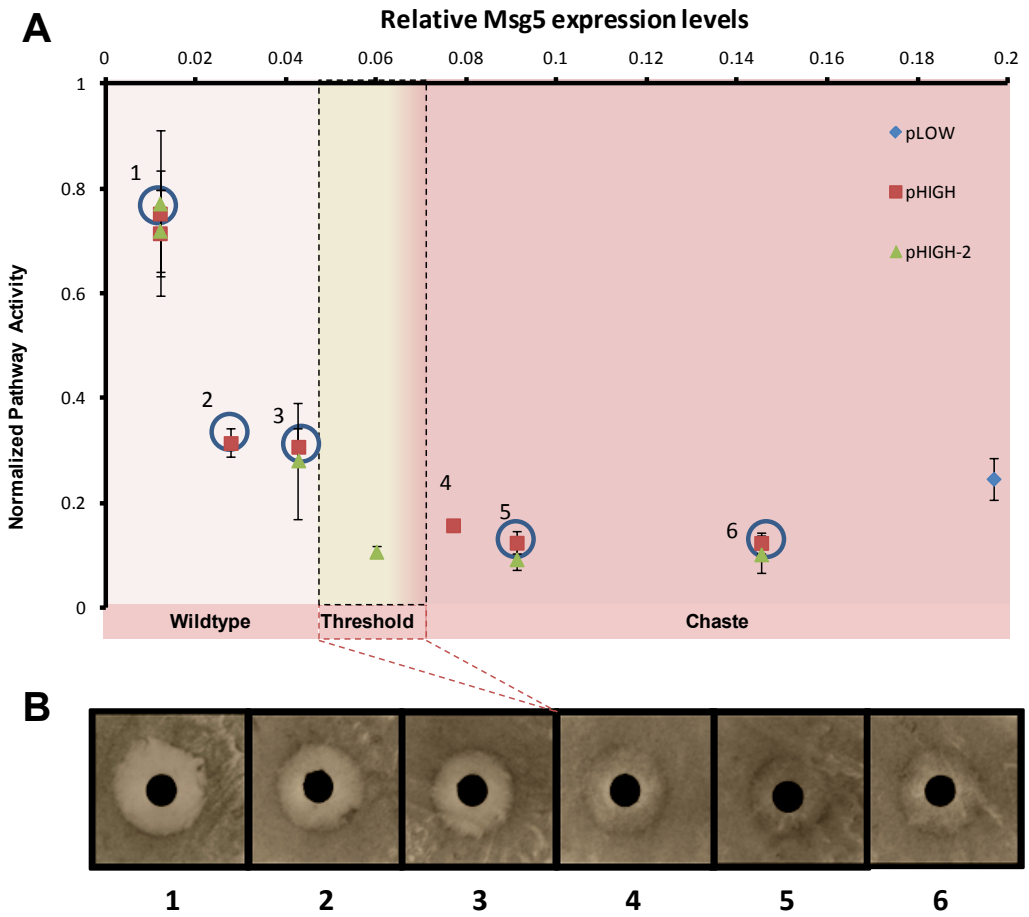
Supplementary Figure 2.7. Constitutive expression of Msg5 routes cells to for entire range of promoter strengths. Within the range of promoter strengths, Msg5 expression is above the threshold required for routing to the chaste fate.



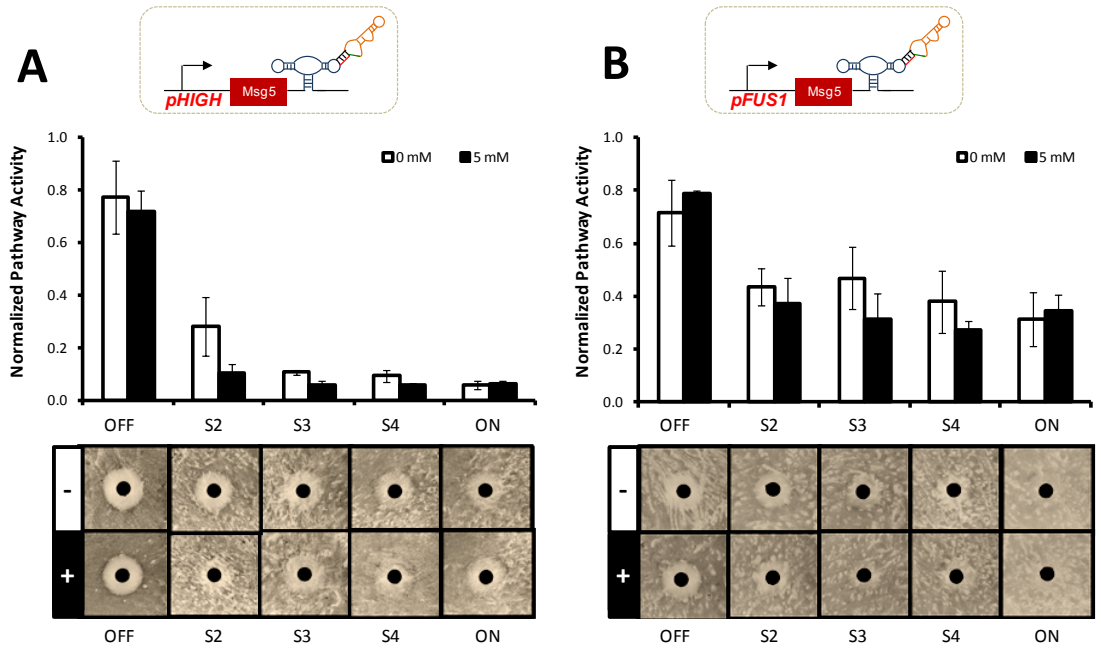
Supplementary Figure 2.8. Histograms show population distribution differences for constitutive and feedback network diverters. A. Histogram for pHIGH Msg5 ON results in a significant increase in the low pFUS1-GFP population relative to OFF. **B.** A low pFUS1-GFP population is minimal in both ON and OFF controls for pFUS1-Msg5. **C.** Histograms evaluating only the high pFUS1-GFP population for pHIGH-Msg5 ON and OFF has populations that mirror pFUS1-Msg5 ON and OFF, respectively, in **D.**



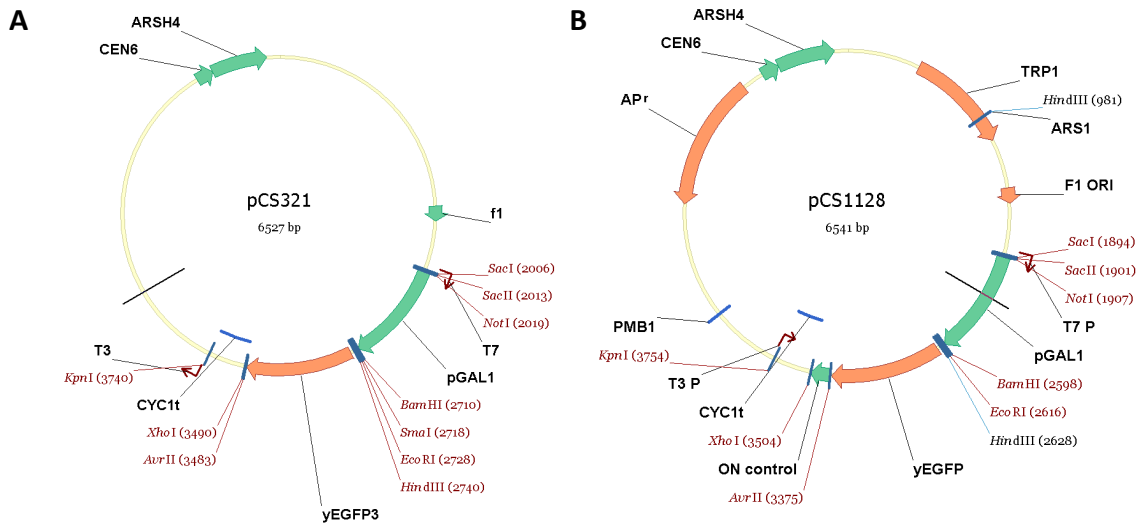
Supplementary Figure 2.9. Neglecting the low pFUS1-GFP population, both feedback architectures show similar mean levels of pathway activity are required for diverting pathway response to the chaste phenotype. Mean values were calculated excluding the low pFUS1-GFP population.



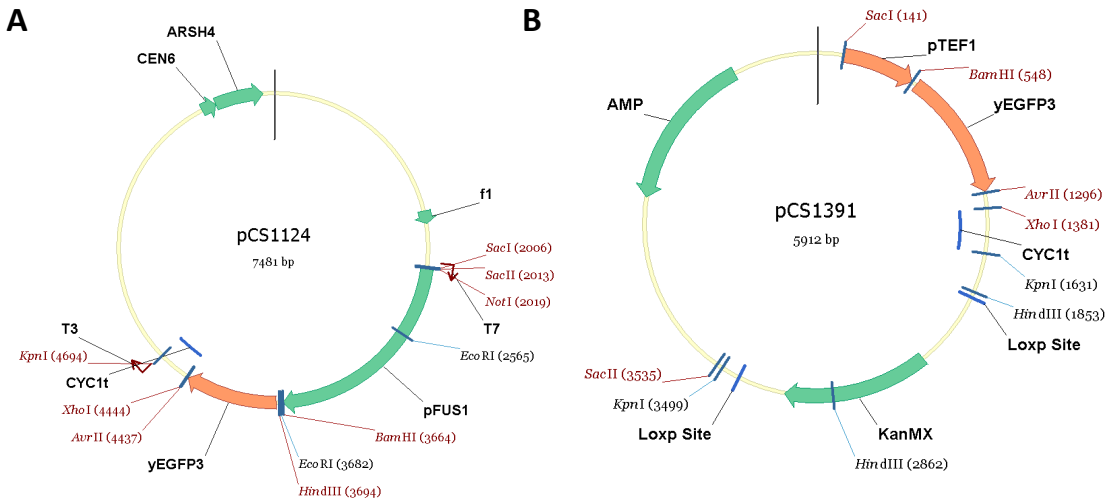
Supplementary Figure 2.10. Tracing the pathway response curve to constitutive Msg5 expression including the low GFP population shows similar transitory range for fate divergence. A. Pathway activity significantly decreases from 1% to 10%. B. Above 8% expression cells adopt the chaste fate. Plate data correspond to points with numbered circles in A selected from the pHIGH-SX data set (Figure 8A). Mean values of pathway activity were calculated including the low pFUS1-GFP population. Relative Msg5 expression was calculated from the relative promoter strengths (Supplementary figure 1) paired with the various switches (Supplementary Figure 2) relative to the pHIGH promoter paired with the ON switch control.



Supplementary Figure 2.10. Theophylline-inducible negative network diverters of different architectures conditionally route cells to the chaste phenotype.



Supplementary Figure 2.11. Original plasmids for construction of molecular network diverter and reporters. A. pCS321 **B.** pCS1128.



Supplementary Figure 2.12. A. pCS1124 **B.** pCS1391

Supplementary tables

Supplementary Table 2.1 GFP plasmids

pCS #	pKG#	Cassette	3'UTR	Marker
1128		pGAL1-YEGFP-CYC1t	sTRSV Ctrl	TRP
321		pGAL1-YEGFP-CYC1t	Empty	URA
1124		pFUS1 -YEGFP-CYC1t	Empty	URA
1585		pTEF1 -YEGFP-CYC1t	Empty	URA
	79	pTEF7 -YEGFP-CYC1t	Empty	URA
	80	pCYC1 -YEGFP-CYC1t	Empty	URA
	78	pADH1 -YEGFP-CYC1t	Empty	URA
	74	pSTL1-yEGFP-CYC1t	Empty	URA

Supplementary Table 2.2 Primer sequences

Insert or PCR Product	Primer Name	Sequence
pTEF7	pTEF7-FWD	5'- AA GAGCTC ATA GCT TCA AAA TGT CTC TAC TCC TTT TT
	pTEF7-REV	5'- AAA GGATCC AAC TTA GAT TAG ATT GCT ATG CTT TCT TTC C
pFUS1	pDS71.FUS1-FWD	5'-TTT GCGGCCGC CCA ATC TCA GAG GCT GAG TCT
	pDS71.pFUS1-REV	5'-TTT GGATCC TTT GAT TTT CAG AAA CTT GAT GGC
pADH1	pADH1-FWD	5'- AA GAGCTC AGC TCG ATA TCC TTT TGT TGT TTC C
	pADH1-REV	5'- AA GGATCC ATT GTA TGC TTG GTA TAG CTT GAA ATA TTG TG
pCYC1	pCYC1-FWD	5'- AA GAGCTC CTC GGT ACC CTA TGG CAT GCA TGT
	pCYC1-REV	5'- AAA GGATCCACGAATTGATCCGGTAATTTAGTGTGTG
pSTL1	pSTL1-FWD	5'-AA GAGCTC GAT TCT GAA ATA CTC CTT TTA CAA CCT TTG C
	pSTL1-REV	5'-AAA GGATCC GGT CTA AAA CTT TCT ATG TTC TAT TTT TC
Ste4	Ste4.k2.FWD	5'- AAA GGATCC A AT TAA TA ATG GCA GCA CAT CAG ATG GAC
	Ste4.REV	5'-AAA CCTAGG CTATTGATAACCTGGAGACCATA
Ste50	Ste50. k2.FWD	5' - AAA GGATCC A AT TAA TA ATG GAG GAC GGT AAA CAG G
	Ste50.REV	5- AAA CCTAGG TTA GAG TCT TCC ACC GGG G
Ste11	Ste11.K2.FWD	5'-AAAAAA GGATCC ATTAAATA ATG GAA CAG ACA CAA ACA GCA GAG
	Ste11.Stop.REV	5'- AAAA CCTAGG TCA AATTATGTGTGCATCCAGCCATGGA
Ste7	Ste7.K2.FWD	5'-AAAAAA GGATCC ATTAAATA ATG TTT CAA CGA AAG ACT TTA CAG AGA AGG
	Ste7.AvrII.REV	5'- AAAAA CCTAGG TCA ATG GGT TGA TCT TTC CGA T
Fus3	Fus3.K2.FWD	5'-AAAAAA GGATCC ATTAAATA ATG CCA AAG AGA ATT GTA TAC AAT ATA TCC AG
	Fus3.A. STOP.REV	5'-AAAAAA CCTAGG CTA ACTA AAT ATT TCG TTC CAA ATG AGT TTC TTG AGG
Msg5	MSG5.K2.FWD	5'- AAA GGATCC A AT TAA TA GTGCACATGCAATTTCAC
	Msg5.REV	5'-AAAA CCTAGG TTAAGGAAGAAACATCATCTG
Hog1	Hog1.FWD	5'-AAA GGATCC ATG ACC ACT AAC GAG GAA TTC ATT AGG A
	Hog1.k2.FWD	5'-AAA GGATCC A AT TAA TA ATG ACC ACT AAC GAG GAA TTC ATT AGG A
	Hog1.REV	5'-AAAA CCTAGG TTACTGTTGGAACTCATTAGCG
Pbs2	Pbs2. k2.FWD	5' - AAA GGATCC A AT TAA TA ATG GAA GAC AAG TTT GCT AAC CTC
	Pbs2.REV	5- AAA CCTAGG CTA TAA ACC ACC CAT ATG TAA TGC CG
Genomic Pbs2 PCR product	Pbs2. Chr.FWD	5' - GCAGATCGAGACGTTAATTTCTC
	Pbs2. Chr.REV	5- TCACGTGCCTGTTTGCTTTT
Ptc1	Ptc1.k2.FWD	5'- AAA GGATCC A AT TAA TA ATG AGT AAT CAT TCT GAA ATC TTA GAA AGG C
	Ptc1.REV	5- AAA CCTAGG TTA GAG GAA GAC AAC CAT GAC C
Genomic Ptc1 PCR product	Ptc1.Chr.FWD	5'- GGCACTGCATTTATCTTTTAAAAATC
	Ptc1. Chr.REV	5- TTGCGCGGTTTATAACGGAT

Supplementary Table 2.3 Galactose-titration plasmids

pCS #	pKG#	Cassette	3'UTR	Marker
1625		pGAL1-Ste4-CYC1t	sTRSV Ctrl	TRP
1483		pGAL1-Ste11-CYC1t	sTRSV Ctrl	TRP
1484		pGAL1-Ste7-CYC1t	sTRSV Ctrl	TRP
1485		pGAL1-Fus3-CYC1t	sTRSV Ctrl	TRP
1486		pGAL1-Msg5-CYC1t	sTRSV Ctrl	TRP
	76	pGAL1-Ste50-CYC1t	sTRSV Ctrl	TRP
	77	pGAL1-Pbs2-CYC1t	sTRSV Ctrl	TRP
	158	pGAL1-Pbs2-CYC1t	sTRSV Ctrl	URA
	157	pGAL1-Hog1-CYC1t	sTRSV Ctrl	TRP
		pGAL1-Ptc1-CYC1t	sTRSV Ctrl	TRP

Supplementary Table 2.4 Primers for mutagenesis

Primer Name	Sequence	Purpose
Ste4. XhoImut.FWD	GTCACTGGTGTGCGATCGAGTCCAGATGG	Remove XhoI in Ste4
Ste4. XhoImut.REV	CCATCTGGACTCGATCGCACACCAGTGAC	
Fus3-XHO.FWD	GGAGAAGATGTTCCCTAGAGTCAACCCGAAAGG	Remove XhoI in Fus3
Fus3-XHO.REV	CCTTTCGGGTTGACTCTAGGGAACATCTTCTCC	
Ste7-AvrII.FWD	GTAACTGGAGAGTTTCCACTAGGTGGGCATAACGA	Remove AvrII in Ste7
Ste7-AvrII.REV	TCGTTATGCCCACCTAGTGGAAACTCTCCAGTTAC	
pCS1441.XhoImut.FWD	ATGCTGGCGGCCGCATCGAGAGATCTAAG	Remove XhoI from disintegrators
pCS1441.XhoImut.REV	CTTAGATCTCTCGATGCGGCCGCCAGCAT	

Supplementary Table 2.5 Switch sequence information

Switches	Alias	Sequence	Basal Expression (0 mM)	Standard Induced Expression (5 mM or 1 mM)
ON	STRSV Ctrl	AAACAAACAAAGCTGTCACCGGATGTGCTTTCCGGTAC GTGAGGTCCGTGAGGACAGAACAGCAAAAAGAAAAAT AAAAACTCGAG	100.0%	100.0%
OFF	sTRSV	AAACAAACAAAGCTGTCACCGGATGTGCTTTCCGGTCTG ATGAGTCCGTGAGGACGAAACAGCAAAAAGAAAAATA AAAACTCGAG	1.2%	1.2%
S1	L2b8-a1 (aka L1N7-1)	AAACAAACAAAGCTGTCACCGGAATCAAGGTCCGGTCT GATGAGTCCGTTGTCCATACCAGCATCGTCTTGATGCCCT TGGCAGGGACGGGACGGAGGACGAAACAGCAAAAAGA AAAATAAAAACTCGAG	3.0%	8.0%
S2	L2b8-a1-t41 (aka L1N7-41)	AAACAAACAAAGCTGTCACCGGAATCAAGGTCCGGTCT GATGAGTCCGTTGCGTATACCAGCATCGTCTTGATGCCCT TGGCAGACGGTAGACGGAGGACGAAACAGCAAAAAGA AAAATAAAAACTCGAG	3.6%	12.7%
S3	L2b8-ta47 (aka S47)	AAACAAACAAAGCTGTCACCGGATGTGCTTTCCGGTCTG ATGAGTCCGTTGAGTATACCAGCATCGTCTTGATGCCCT TGGCAGACTGTATACGGAGGACGAAACAGCAAAAAGA AAAATAAAAA	6.1%	33.6%
S4	L2b8 (aka L2b12)	AAACAAACAAAGCTGTCACCGGATGTGCTTTCCGGTCTG ATGAGTCCGTTGTCCATACCAGCATCGTCTTGATGCCCTT GGCAGGGACGGGACGGAGGACGAAACAGCAAAAAGAA AAATAAAAACTCGAG	9.1%	36.7%
S5	L2b1	AAACAAACAAAGCTGTCACCGGATGTGCTTTCCGGTCTG ATGAGTCCGTGTCCATACCAGCATCGTCTTGATGCCCTTG GCAGGGACGGGACGAGGACGAAACAGCAAAAAGAAAA ATAAAAA	39.8%	75.3%
S1tc	L2b12tc-11	AAACAAACAAAGCTGTCACCGGATGTGCTTTCCGGTCTG ATGAGTCCGTTGTTGAAAACATACCAGATTTCGATCTGG AGAGGTGAAGAATTCGACCACCTCATTTCAACGGAGGA CGAAACAGCAAAAAGAAAAATAAAAACTCGAG	3.7%	4.8%
S2tc	L2b8tc-a1 (aka L2b12tc-1)	AAACAAACAAAGCTGTCACCGGAATCAAGGTCCGGTCT GATGAGTCCGTTGTCCAAAACATACCAGATTTCGATCTG GAGAGGTGAAGAATTCGACCACCTGGACGGGACGGAGG ACGAAACAGCAAAAAGAAAAATAAAAACTCGAGCC	3.0%	10.8%
S4tc	L2b8tc (aka L2b12tc)	AAACAAACAAAGCTGTCACCGGATGTGCTTTCCGGTCTG ATGAGTCCGTTGTCCAAAACATACCAGATTTCGATCTGG AGAGGTGAAGAATTCGACCACCTGGACGGGACGGAGGA CGAAACAGCAAAAAGAAAAATAAAAACTCGAG	8.4%	43.9%
S3tc	L2b12tc-NheI-L2bOFF1 (aka Stc-OFF1)	AAACAAACAAAGCTGTCACCGGATGTGCTTTCCGGTCTG ATGAGTCCGTTGTCCAAAACATACCAGATTTCGATCTGG AGAGGTGAAGAATTCGACCACCTGGACGGGACGGAGGA CGAAACAGCAAAAAGAAAAATAAAAAGCTAGGAAAC AAACAAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGA GTCCGTGTTGCTGATACCAGCATCGTCTTGATGCCCTTGG CAGCAGTGGACGAGGACGAAACAGCAAAAAGAAAAAT AAAAA	4.2%	21.2%

Supplementary Table 2.6 Msg5 plasmids

pX-Msg5-CYC1t			
pCS #	**pKG#**	**Promoter**	**3'UTR**
	88	pADH1	Empty
	145	pCYC1	Empty
	89	pTEF7	Empty
1487		pFUS1	Empty
	200	pADH1	ON
	154	pCYC1	ON
	123	pTEF7	ON
	121	pFUS1	ON
	199	pTEF1	ON
pFUS1-Msg5-CYC1t			
pCS #	**pKG#**	**Parent**	**3'UTR**
	121	pCS1487	ON
	45	pCS1487	S5
	48	pCS1487	S4
	286	pCS1487	S3
	220	pCS1487	S2
	168	pCS1487	S1
	137	pCS1487	OFF
pTEF7-Msg5-CYC1t			
pCS #	**pKG#**	**Parent**	**3'UTR**
	123	pKG89	ON
	112	pKG89	S5
	122	pKG89	S4
	293	pKG89	S3
	219	pKG89	S2
	169	pKG89	S1
	136	pKG89	OFF

Supplementary Table 2.7 Ste4 plasmids

pX-Ste4-CYC1t		
pKG#	**Promoter**	**3'UTR**
90	pADH1	Empty
87	pTEF7	Empty
84	pFUS1	Empty
159	pADH1	ON
203	pCYC1	ON
144	pTEF7	ON
173	pFUS1	ON
201	pTEF1	ON
pFUS1-Ste4-CYC1t		
pKG#	**Parent**	**3'UTR**
228	pKG227	ON
234	pKG227	S4tc
298	pKG227	S3tc
227	pCS2094	S2tc
274	pKG227	S1tc
233	pKG227	OFF
pFUS1-Ste4-CYC1t		
pKG#	**Parent**	**3'UTR**
173	pKG84	ON
215	pKG84	S5
64	pKG84	S4
161	pKG84	S2
176	pKG84	S1
124	pKG84	OFF
95	pKG84	S5tc
189	pKG84	S4tc
179	pKG84	S2tc
180	pKG84	S1tc
pTEF7-Ste4-CYC1t		
pKG#	**Parent**	**3'UTR**
144	pKG87	**ON**
272	pKG87	S4tc
296	pKG87	S3tc
177	pKG87	S2tc
178	pKG87	S1tc
152	pKG87	OFF
213	pKG87	S5tc
214	pKG87	S4tc
153	pKG87	S2tc
198	pKG87	S1tc
pADH1-Ste4-CYC1t		
pKG#	**Parent**	**3'UTR**
159	pKG90	ON
86	pKG90	S5
99	pKG90	S4
172	pKG90	S2
171	pKG90	S1
160	pKG90	OFF
pTEF7-Ste4-CYC1t		
pKG#	**Parent**	**3'UTR**
321	pCS1128	S4tc
295	pCS1128	S3tc
188	pCS1128	S2tc
182	pCS1128	S1tc
181	pCS1128	OFF

Supplementary Table 2.8 Plasmids for strain construction

pCS #	pKG#	Cassette	Integration Loci	Marker
1543		pCS1439 with XhoI removed; disintegrator [37]	FCY1	URA
	98	pCS1543 + pSTL1 -YEGFP-CYC1t	FCY1	URA
1391		LoxP Integrating plasmid [35]	Primer-specific	G418
	100	pCS1391+pFUS1-yEGFP-CYC1t	Primer-specific	G418

Supplementary Table 2.9 Yeast strains

CSY#	KGY#	Description
364		EY1119 from [36]; W303a Δsst1 Δkss1::HIS3
408		CSY364 Δgal2::KanR
532		CSY364 Δgal2::pFUS1-GFP-CYC1t-loxP-KanR
840	144	CSY364 Δtrp1::pFUS1-yEGFP3-CYC1t-loxP-KanR
	145	CSY364 Δgal21::pFUS1-yEGFP3-CYC1t-loxP-KanR
	139	CSY408 FCY1::pSTL1-GFP-CYC1t

Supplementary Table 2.10 Primers for integration

Loci	Sequence	Direction
GAL2	ATGCACCTTATTCAATTATCATCAAGAATAGTAATAGTTAAGTAAACACAAGATTAACATAGATTGTACTGAGAGTGCAC	FWD
	ATGATAATTAAAATGAAGAAAAAACGTCAGTCATGAAAAATTAAGAGAGATGATGGAGCGCGACTCACTATAGGGAGACC	REV
TRP	GTATACGTGATTAAGCACACAAAGGCAGCTTGGAGTATGTCTGTTATTAATTTCACAGGAAGATTGTACTGAGAGTGCAC	FWD
	TTGCTTTTCAAAAGGCCTGCAGGCAAGTGCACAAACAATACTTAAATAAATACTACTCAGCGACTCACTATAGGGAGACC	REV

Chapter 3

Constructing synthetic gene networks to control decision-making in the yeast mating pathway

Abstract

Cells utilize internal molecular networks to direct cell fate via the integration and processing of extracellular information. Control of these decision-making pathways offers the potential to intervene in aberrantly activated programs as well as direct complex, multicellular tasks such as tissue development. Programming of diverse cellular behaviors may be accomplished by the development of synthetic circuits capable of activating and attenuating the response from internal signaling pathways. Previously constructed positive and negative molecular network diverters encoding a single pathway regulator demonstrated small-molecule-dependent routing to divergent alternative fates. However, simultaneous integration of these previously demonstrated positive and negative diverters results in diverter antagonism and fails to allow dual-routing to both alternative fates. In this work, we demonstrate the construction of more complex diverter networks with integrated positive and negative routing functions that allow the conditional induction of alternative cell fates based on small-molecule input. By constructing networks with two differentially regulated expression modules composed of stringent RNA-based controllers we were able to limit diverter antagonism, amplify pathway activation, and induce pathway attenuation allowing genetically identical cells to be conditionally routed to one of three fates in response to environmental cues. In constructing these networks we identified sensitive parameters for balancing and tuning diverter function and demonstrate the rational tuning of both diverters to allow dual-fate routing. Further, we demonstrate the construction of networks that suppress subthreshold noise in gene expression to robustly amplify differences in environmental input and achieve enhanced resolution between triggered and non-triggered cell populations.

Expanding synthetic control to mammalian systems via the elucidated molecular network diverter strategies and principles will facilitate the construction of sophisticated synthetic programs to achieve higher-order cellular functions such as the patterning of cell fate, tissue homeostasis, and autonomous immune surveillance.

Introduction

Biological systems process environmental signals via native regulatory networks directing complex processes such as development, tissue homeostasis, and the immune system response [1]. The temporal and spatial distribution of molecular signals enables responses to be coordinated in time and space via the distributed processing of these signals by individual cells, resulting in orchestrated system responses such as organogenesis, wound healing, and pathogen elimination [2-4]. Synthetic circuits that can interface with and redirect these processes offer the potential to orthogonally regulate complex cellular behaviors. However, construction of an interface that provides for specific pathway information to be translated via synthetic circuitry into changes in the native network still poses a major challenge in synthetically regulating cell fate [5, 6]. Further, synthetic circuits that can be tuned via their component parts as well as architecture will facilitate expansion of such circuits to native networks with a range of associated properties [7]. The development of parts that allow orthogonal tuning of system performance without modification of other potentially constrained regulatory elements (e.g., promoters, genes) will enable greater flexibility in adapting synthetic circuits to be regulators of cellular behavior.

In previous work, a type of synthetic gene circuit, referred to as a molecular network diverter, was applied to controlling signaling and fate decisions in a model MAPK pathway in response to distinct environmental triggers. Specifically, we constructed positive and negative diverters that were used to turn the activity of the *Saccharomyces cerevisiae* mating pathway on and off, respectively. The positive network diverter activated the pathway in the absence of the canonical pathway input, thereby

routing cells to the alternative “promiscuous” fate. The negative network diverter inactivated the pathway even in the presence of the canonical pathway input (pheromone), routing cells to the “chaste” fate. The positive and negative diverters are programmed to respond to distinct small-molecule inputs, allowing independent channels by which to activate each diverter. Thus, we propose integrating positive and negative diverters within the same cell to achieve conditional routing of cells to three distinct cell fates. In the absence of either environmental trigger, cells will adopt the wild-type fate as the diverters run quiescently. However, when called by their respective triggers, each diverter is intended to route cells to the programmed alternative fates: chaste or promiscuous.

Molecular network diverters are composed from expression modules containing three genetic parts; a promoter, a pathway regulator, and a RNA-based controller (Figure 3.1). The properties of these parts dictate the function and performance of the diverter. The sign of the diverters, positive or negative, is determined by whether the pathway regulator activates or attenuates pathway activity in the network, respectively. The promoter determines the mode of expression, feedback or nonfeedback, as well as the strength of the diverter activity. RNA controllers provide a second layer of control by which to tune diverter activity and allow the molecular network diverter to respond to environmental signals. Small-molecule-responsive RNA-based controllers, also called RNA switches, are an engineered class of non-coding RNA that allows for conditional control from any promoter-gene pair via a chemical trigger [8]. Choice of the switch specifies the small-molecule trigger that activates the diverter and tunes the diverter activity. Combining these three parts together in various configurations generates a

variety of expression modules. Dependent on the context of implementation, these expression modules can perform independently as molecular network diverters or be integrated into larger networks of synthetic circuitry. Previous work has shown that single-module diverters, composed of one expression module, can be implemented to conditionally route cells to an alternative fate. However, to access both alternative fates the simultaneous integration of positive and negative diverters is required.

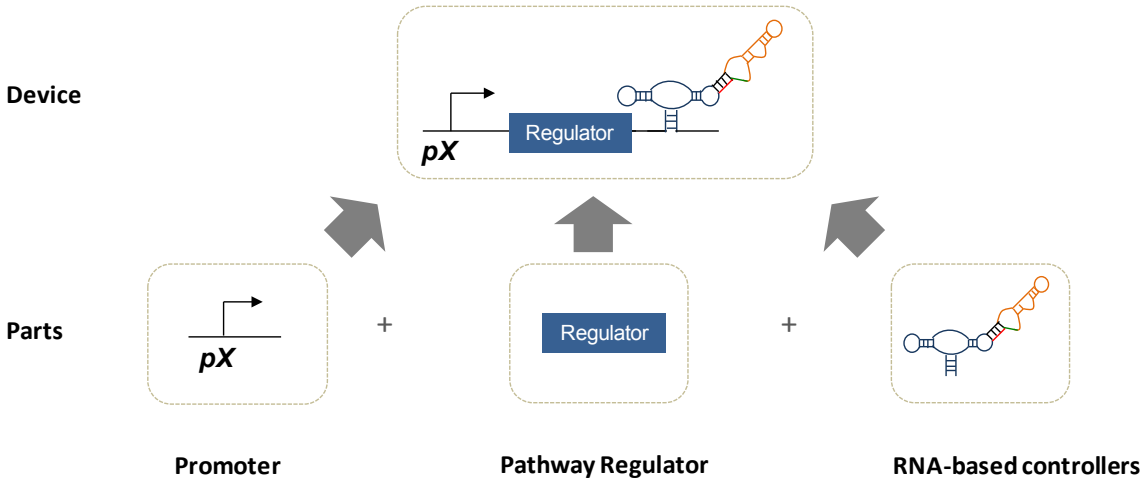


Figure 3.1. Composition of an expression module from well-defined parts. Molecular network diverters contain various expression modules which are composed of a promoter, a pathway regulator, and a RNA-based controller, also called a switch. The promoter specifies the expression mode (e.g., constitutive or feedback) which determines the network connectivity. The promoter and switches combine to determine the expression strength and thus the activity of the diverter. Additionally, switches specify which small-molecule input regulates expression and the range of expression across ligand concentrations. The pathway regulator's interaction with the native molecular network determines the pathway response curve. The pathway response curve can be altered by changes in network connectivity due to feedback expression of regulators.

One potential challenge that may be encountered in integrating network diverters encoding opposing functions is antagonization, or competition between the opposing activities encoded in the diverters. Antagonization may hinder the diverter's ability to route cells to both alternative fates. Specifically, the positive and negative diverters measurably change the basal activity of the pathway in the absence of their respective triggers. While these sub-threshold changes in activity are insufficient to route cells to an

alternative fate, they may antagonize the activity and routing capability of an opposing diverter. Issues with basal expression levels from synthetic regulatory networks impinging on circuit performance have been previously reported [9-11]. One strategy used to increase the stringency associated with a synthetic network is to incorporate additional layers of post-transcriptional control via noncoding RNA control elements [5, 12, 13]. While strategies for reducing basal expression may also reduce triggered expression levels, modification of the regulatory network architecture through the incorporation of feedback loops may supplement the reduction in triggered expression levels via the amplification of the pathway response to the trigger.

Natural biological systems utilize feedback architectures to ensure proper function of diverse cellular processes. For example, positive and negative feedback loops are used to facilitate robustness in developmental processes by introducing ultrasensitivity and by maintaining homeostasis, respectively, [14-16]. More complex control architectures, such as coherent and incoherent feedfoward loops, have been shown to function as persistence detectors, pulse generators, and response accelerators in natural and synthetic systems [17-19]. The proper configuration of such regulatory architectures within developmental pathways is posited to contribute to the robustness of multicellular organization critical for higher eukaryotes [14, 20-22]. For example, modeling of cell fate determination in flowers has predicted that network architecture in these systems constrains pathway activity to a few stable states independent of the chosen initial conditions or model parameters [23]. In addition, synthetic circuits that reshape natural molecular pathways have shown that cell fate can be routed by altering the native network topology [24]. Thus, network topology itself can ensure robust adoption of particular cell fates.

In this chapter, we show that simultaneously integrating positive and negative diverters in the yeast mating pathway, optimized for individual routing functions, fails to achieve routing to both alternative fates. In particular, routing to the promiscuous fate is hampered across the range of activities exhibited by the positive feedback diverters when paired with any negative diverter. Resistance introduced by basal level expression of Msg5 from the negative (resistance) diverter may prevent pathway activation above the requisite threshold to trigger positive feedback-induced amplification. The addition of a positive expression (booster) module to the positive feedback diverter effectively counteracts this network antagonism to construct the amplifying diverter (Figure 3.2). The amplifying diverter allows promiscuous routing in the presence of the resistance diverter. However, the resistance diverter is unable to route cells to the chaste fate when implemented with the amplifying diverter. The data indicate that at high levels of pheromone, conditions in which positive feedback is amplified, the amplifying diverter overwhelms pathway inhibition by the resistance module. The addition of a negative feedback module to the resistance diverter results in an attenuating diverter that effectively balanced pathway response. Utilizing this double-module strategy, we successfully integrated the amplifying and attenuating diverters to achieve dual-fate routing to both alternative programmed cell fates. Our work shows that the performance of the amplifying and the attenuating diverters is highly sensitive to the strengths of the positive feedback and resistance modules. Further, our findings have important implications for the construction of networks that suppress noise amplification, amplify differences in environmental stimuli, and trigger robust cellular phenotypic programs in the face of antagonistic signals.

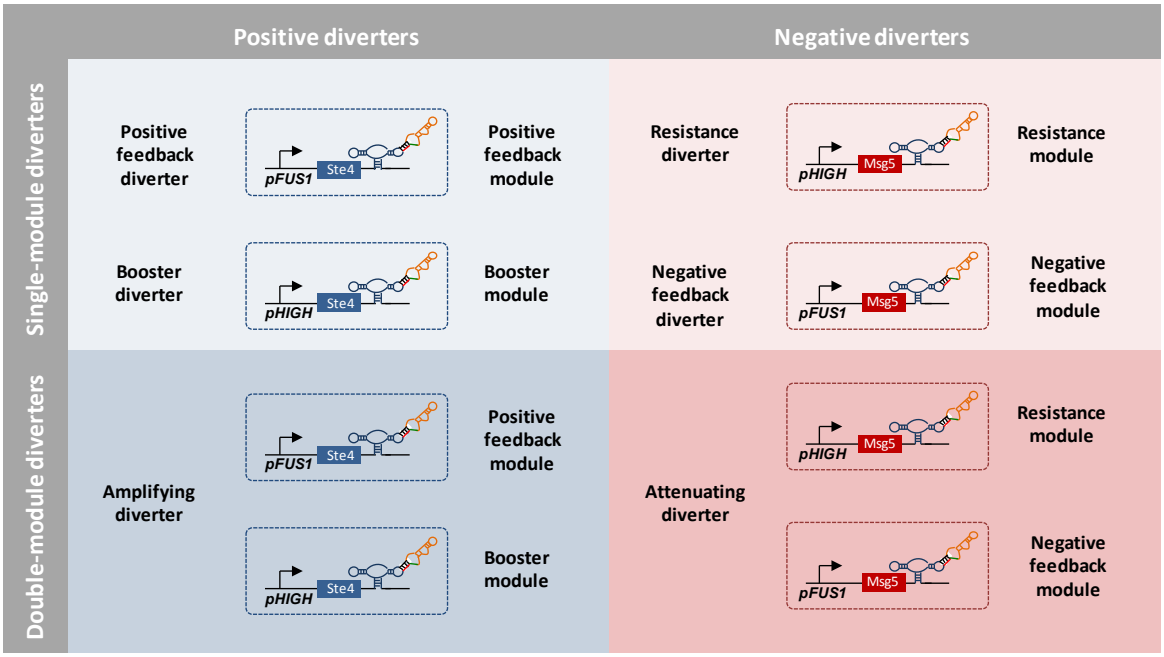


Figure 3.2. Molecular network diverters are composed from single- and double-expression modules that allow fate-routing dependent on small-molecule input. Six different types of diverters were constructed. The positive feedback diverter, the booster diverter, and the amplifying diverter represent the three types of positive diverters. The resistance diverter, the negative feedback diverter, and the attenuating diverter represent the three types of negative diverters. Diverter performance is determined by the composition of their expression modules. Four of the diverter types are single-module diverters containing only one expression module. The amplifying and the attenuating diverters are double-module diverters. The amplifying diverter is constructed with a positive feedback module and booster module. The attenuating diverter contains a resistance module and a negative feedback module.

Results

Simultaneous expression of single-module positive and negative diverters fails to route cells to either fate

The original molecular network diverters were optimized to independently route fate decisions in the yeast mating pathway and exhibited different regulatory architectures. The positive diverter integrated a positive network regulator (Ste4) with a feedback promoter (pFUS1) and a tetracycline-responsive RNA switch to compose the positive feedback diverter which activates pheromone-independent signaling through the

pathway in the presence the tetracycline, resulting in the so-called promiscuous fate. The negative diverter integrated a negative network regulator (Msg5) with a constitutive promoter (pHIGH) and a theophylline-responsive RNA switch constructing the resistance diverter to inhibit signaling through the pathway in the presence of theophylline even in the presence of pheromone, resulting in the so-called chaste fate.

We examined an initial dual diverter architecture based on integrating the positive and negative diverters optimized for independent cell fate routing. To examine optimal pairing of diverters encoding antagonistic functions, we paired the positive feedback diverter (pFUS1-Ste4-S3tc) with diverse resistance diverters incorporating a set of RNA switches exhibiting a range of activities (pHIGH-Msg5-Sx) (Supplementary Figure 3.1). Pathway activity was determined by measuring GFP expression levels from a transcriptional fusion construct (*pFUS1*-GFP, GFP fused to a mating responsive promoter). Fate routing was determined by observing mating-associated cell cycle arrest via halo assays (Materials and Methods). Cells adopting the chaste fate exhibit reduced halo formation as cells resist pheromone-induced cell cycle arrest. Cells adopting the promiscuous fate demonstrate persistent, pheromone-independent cell-cycle arrest outside the canonical halo region as indicated by reduced cell growth across the entire plate. The results indicate that within the dual diverter architecture the two diverters antagonize one another, diminishing diverter performance and leading to less robust fate switching (Figure 3.3). While several dual diverter configurations preserve weak routing to the chaste fate, promiscuous routing was not observed from any of the dual diverter configurations. In addition, pathway activation was not improved by increasing the strength of the positive feedback module (Supplementary Figure 3.2). The data indicate

that switching to the promiscuous fate may be hindered by basal expression levels of Msg5 from the resistance module. Resistance may inhibit superthreshold pathway activation, which is necessary for amplification via the positive feedback loop, resulting in low pathway activation for a significant fraction of cells even in the presence of the positive molecular trigger (tetracycline).

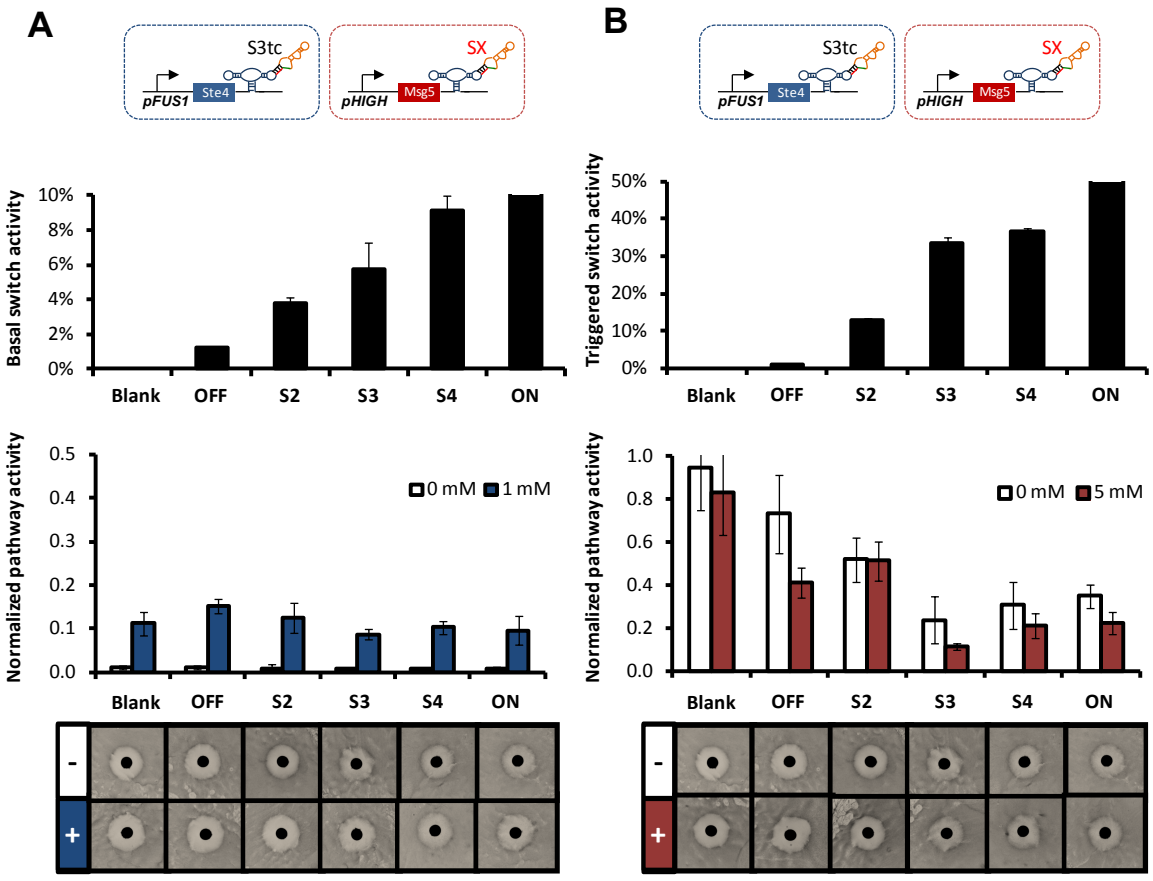


Figure 3.3. Integration of positive and negative single-module diverters fails to achieve dual-fate routing. A. A single-module positive feedback diverter, pFUS1-Ste4-S3tc, fails to route to the promiscuous fate in the presence of resistance diverters incorporating a range of switch strengths. The Blank control bears a plasmid without either diverter. Increasing positive feedback strength does not improve pathway activation from the positive diverter (Supplementary Figure 3.1). **B.** Negative diverters incorporating constitutive expression from pHIGH of Msg5 and switches of varying strength show significant reduction of pathway activity, but weak routing to the chaste fate in halo assays.

Building network architectures to amplify trigger-induced switching to activate the mating pathway in the presence of antagonistic signals

We explored alternate architectures for the positive diverter to support increased pathway activation. A number of strategies can be implemented for increasing the activity of the positive feedback diverter, including modifying the switch or promoter in the positive feedback module to exhibit increased activity, or incorporating an additional copy of the positive feedback module. However, such modifications are expected to raise the basal levels of Ste4 at high pheromone input, potentially inhibiting the performance of the negative diverter. As an alternative strategy, we added a second module encoding constitutive expression of Ste4, or a booster module, to the positive feedback diverter to construct the amplifying diverter. Addition of the booster module has two notable advantages over a second feedback module. First, the booster module is insensitive to the resistance imposed by the negative diverter, such that Ste4 expression from this module is independent of the negative diverter's effect on pathway activity. Second, basal expression of Ste4 from the booster module is constant across the range of pheromone input, such that antagonization of the negative diverter does not increase from this module in the absence of its trigger molecule within the high-pheromone input regime. Thus, the incorporation of the booster module into the positive diverter architecture offers the potential to enhance the performance of the positive diverter, while imposing minimal effect on the performance of the negative diverter.

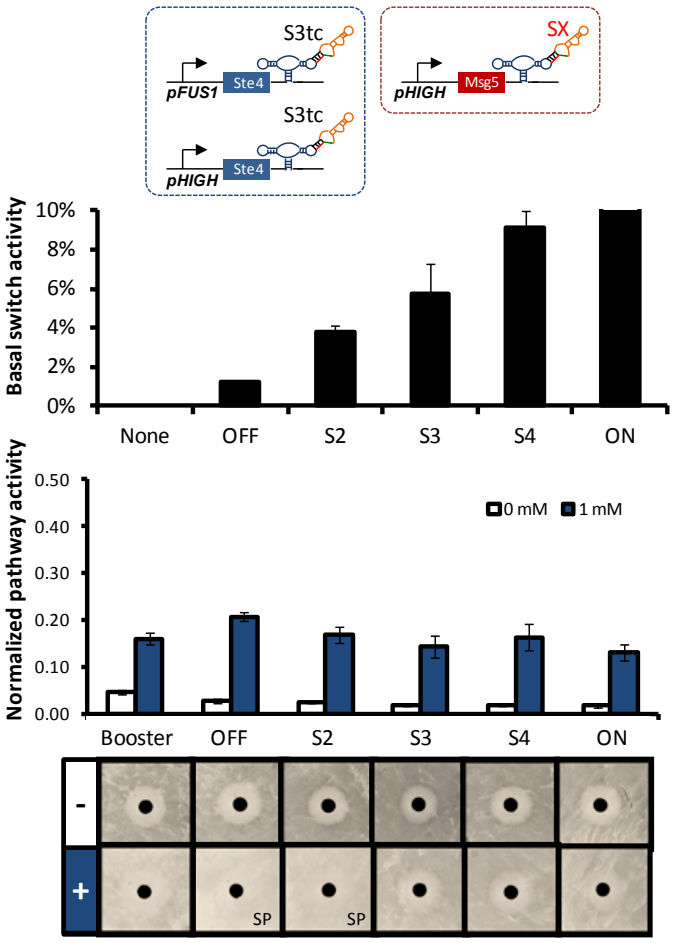


Figure 3.4. Addition of booster module to positive feedback diverter enhances fate switching in the presence of the resistance diverter. An amplifying diverter with S3tc regulating both expression modules is paired with resistance diverters incorporating pHIGH-Msg5 expression with switches of various activities. Pathway activation via this amplifying diverter diminishes as the basal activity of the resistance module increases. Networks constructed with this amplifying diverter and low Msg5 resistance (OFF, S2) show enhanced fate switching in halo assays relative to the booster diverter. However, for resistance modules with switch basal levels above S2 small-molecule triggered fate routing is weaker. "SP" denotes strong promiscuous routing in the plate assays.

The ability of the amplifying diverter to overcome antagonism from the negative diverter and achieve programmed routing to the promiscuous fate was examined. The amplifying diverter was composed of a tetracycline-responsive RNA switch regulating both expression modules (feedback: pFUS1-Ste4-S3tc; booster: pHIGH-Ste4-S3tc) and paired with resistance diverters incorporating a set of theophylline-responsive RNA switches exhibiting a range of activities (pHIGH-Msg5-Sx). Pathway activation and fate routing were evaluated as described previously. Results demonstrate that pathway

activation through the amplifying diverter diminishes as the basal activity of the resistance diverter increases (Figure 3.4). Network configurations with the amplifying diverter and low-strength resistance modules (i.e., OFF, S2) exhibit enhanced fate switching in halo assays relative to the booster module alone. Halo assays indicate that small-molecule-triggered routing to the promiscuous fate is noticeably weaker for resistance modules exhibiting greater activity. The data support that in the absence of the booster module, basal pathway activity in the presence of a resistance module is insufficient for amplification of pathway activity via the positive autoregulatory loop even in the presence of the molecular trigger (tetracycline). Thus, we postulate that increased expression from the booster module initiates increased pathway activity. Once the pathway activity crosses the requisite threshold, Ste4 expression is amplified and reinforced by the positive feedback module.

The pathway activation ratio (PAR) in the absence and presence of the environmental trigger provides a metric for evaluating a diverter's ability to facilitate cell fate reprogramming. A large PAR value is generally desirable for regulatory networks that route to divergent fates as it indicates a greater difference between the pathway activities of triggered and non-triggered cells. However, while the PAR value offers one performance metric for evaluating molecular network diverters, it is not a sufficient metric to determine a diverter's ability to route cells to an alternative fate. Diverters may achieve moderate PAR values by modifying basal pathway activity, but fail in routing if the triggered pathway activity does not cross the threshold for fate routing. Conversely, diverters exhibiting modest PAR values may effectively route fate provided they are

configured to cross the threshold of fate divergence. An optimal diverter configuration achieves both routing capability and a larger PAR value.

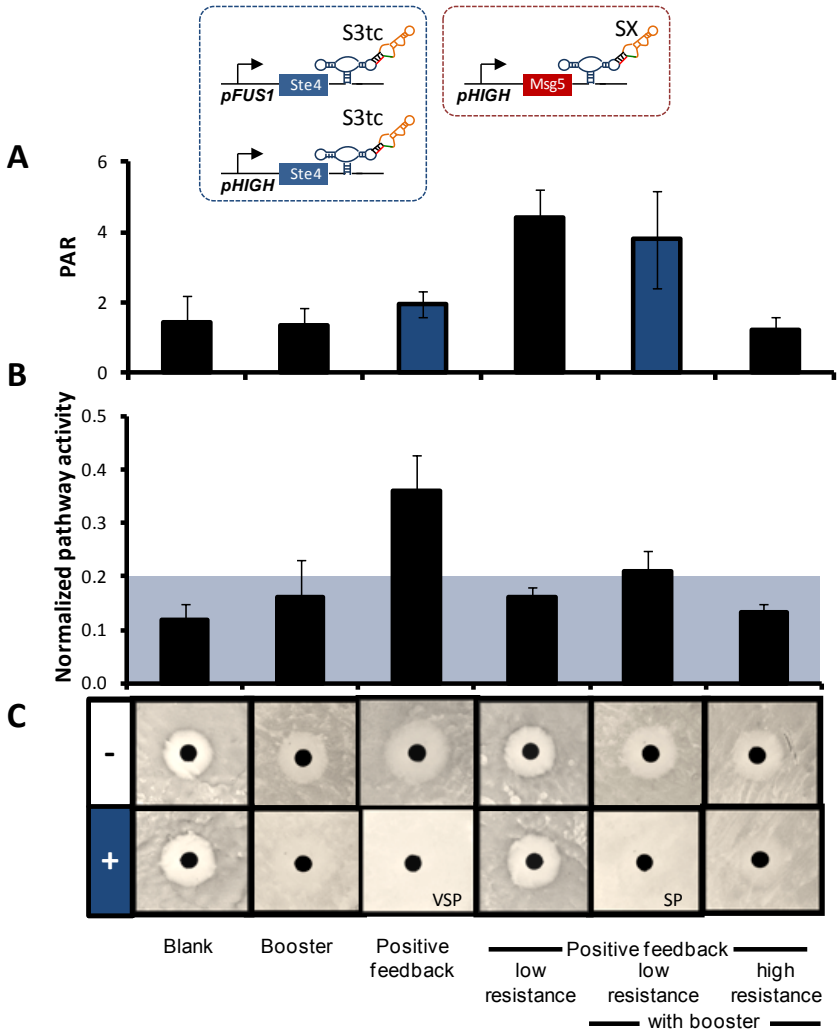


Figure 3.5. Pathway activation ratio is enhanced for positive feedback diverters in network configurations including a low-strength resistance module. A. Addition of a low strength resistance module to the synthetic network increases the PAR value of the positive feedback diverters with and without the booster module by lowering basal level expression. **B.** Pathway activity data for tetracycline triggered cells over a range of network configurations indicate that only two configurations sufficiently increase pathway activity above the requisite fate switching threshold. **C.** Halo assays demonstrate that only networks configured with the positive feedback modules (positive feedback diverter and amplifying diverter paired low resistance module) strongly route to the promiscuous fate. "SP" and "VSP" denote strong and very strong promiscuous routing in the plate assays, respectively.

The PAR values for positive diverters incorporating a variety of network configurations were determined by measuring the ratio of pathway activity in the presence and absence of the environmental trigger (tetracycline). Despite activating the

pathway sufficiently to route cells to the promiscuous fate, positive feedback diverter exhibits a lower PAR value due to the increased basal activity of the pathway associated with this diverter (Figure 3.5A). Further, the high basal pathway activity associated with the positive feedback diverter provides little resolution between the triggered and non-triggered populations (Figure 3.6A). The implementation of a positive feedback diverter with a low activity resistance diverter reduces the basal pathway activity, thereby increasing the PAR value for this network configuration and increasing population separation. However, the resistance module also suppresses amplification of pathway activity via the autoregulatory loop, thus inhibiting routing to the promiscuous fate (Figure 3.6B, Figure 3.5). Robust routing to the promiscuous fate is only achieved when a low activity resistance diverter is paired with an amplifying diverter (Figure 3.5B, C), resulting in enhanced separation between the triggered and non-triggered populations (Figure 3.6C). Altering the network configuration by incorporation of a high activity resistance diverter reduces the pathway activity, resulting in a diminished PAR value and weaker routing to the promiscuous fate. We postulate that the resistance module enhances differentiation between triggered and non-triggered populations by buffering the effect of subthreshold variations in Ste4 levels on pathway activity, reducing the amplification of noise in the system. Further, the strength of the resistance module modulates the triggered output from the amplifying diverter and thus tunes routing to the promiscuous fate. Therefore, the ability to rationally tune module and diverter activities is critical for configuring network architectures that can cross thresholds of fate divergence in the presence of environmental signals. Optimal diverter performance is enhanced by

networks that suppress noise amplification and other subthreshold variations in gene expression to robustly amplify the appropriate input.

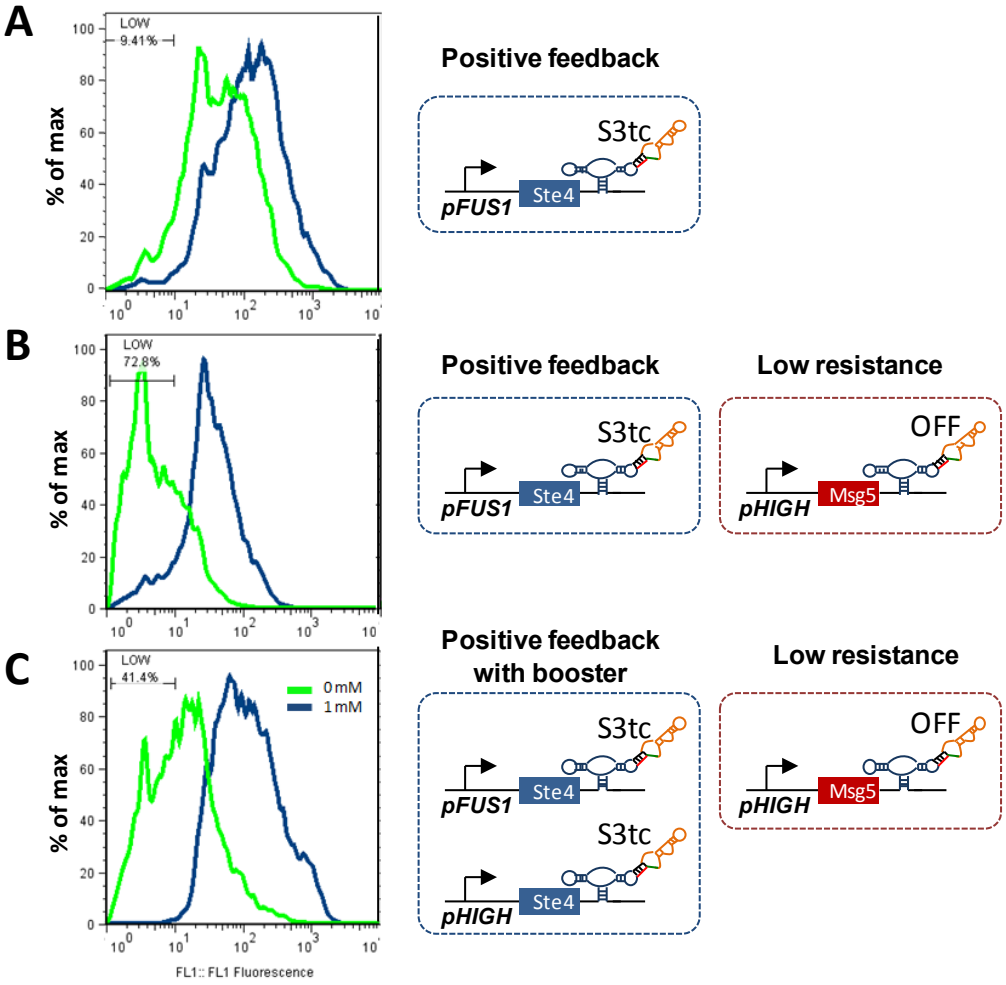


Figure 3.6. Structuring networks to amplify switching by layering positive feedback with a resistance module and a booster module. A. A histogram of pFUS1-GFP for the positive feedback diverter pFUS1-Ste4-S3tc shows significant overlap between the populations in the presence and in the absence of tetracycline, resulting in a low PAR value (1.9). Despite a low PAR value pFUS1-Ste4-S3tc effectively routes cells to the promiscuous fate (Figure 3.4) **B.** Adding the low resistance module to the network configuration via pHIGH-Msg5-OFF decreases overlap between the two populations, increasing the PAR value (4.4). Yet, this network provides insufficient pathway activation to route to the promiscuous fate when trigger with tetracycline (Figure 3.4). **C.** Addition of a booster module, pHIGH-Ste4-S3tc, to the positive diverter constructs the amplifying diverter. The booster module in the amplifying diverter boosts pathway activity, routing cells to the promiscuous fate when triggered. Further, population separation is maintained in this configuration as indicated by the high PAR value (3.8) (Figure 3.4).

Network configurations that balance positive feedback with negative feedback enhance pathway attenuation

While addition of the booster module enhances the ability of the positive diverter to route cells to the promiscuous fate, the increased basal expression from this amplifying diverter diminishes the ability of the resistance diverter to route cells to the chaste fate. We postulate that high levels of pheromone leads to amplification of the positive feedback loop, such that the positive diverter overwhelms the attenuation provided by the resistance diverter. Therefore, we examined the impact of adding a negative feedback module to this network to balance the pathway response.

To counteract the amplification of pathway activity via the positive feedback module in the amplifying diverter, we constructed an attenuating diverter, which incorporates a resistance module and a negative feedback module. An amplifying diverter was configured with a booster module that incorporated a strong tetracycline RNA switch (pHIGH-Ste4-S4tc) and a positive feedback module that incorporated a moderate tetracycline RNA switch (pFUS1-Ste4-S3tc). This amplifying diverter was paired with various attenuating diverters composed of a resistance module (pHIGH-Msg5-S3) and negative feedback modules incorporating a set of theophylline-responsive RNA switches exhibiting a range of activities (pFUS1-Msg5-Sx). Pathway activation and fate routing were evaluated as described previously. The results indicate that pathway attenuation is relatively insensitive to the strength of negative feedback module (Figure 3.7A). While GFP levels exhibit moderate changes with increasing negative feedback strength, halo assays indicate that these diverters robustly route cells to the chaste fate in the presence of theophylline above a threshold of negative feedback strength (S3, S4). In the absence

of theophylline, the network maintains normal halo formation for the negative feedback module regulated via S3. However, the network composed with the negative feedback module regulated by S4 exhibits reduced halo formation even in the absence of theophylline. The data indicate that there exists a threshold of negative feedback strength residing between the basal expression levels of S3 and S4 at which wild-type and chaste fates diverge.

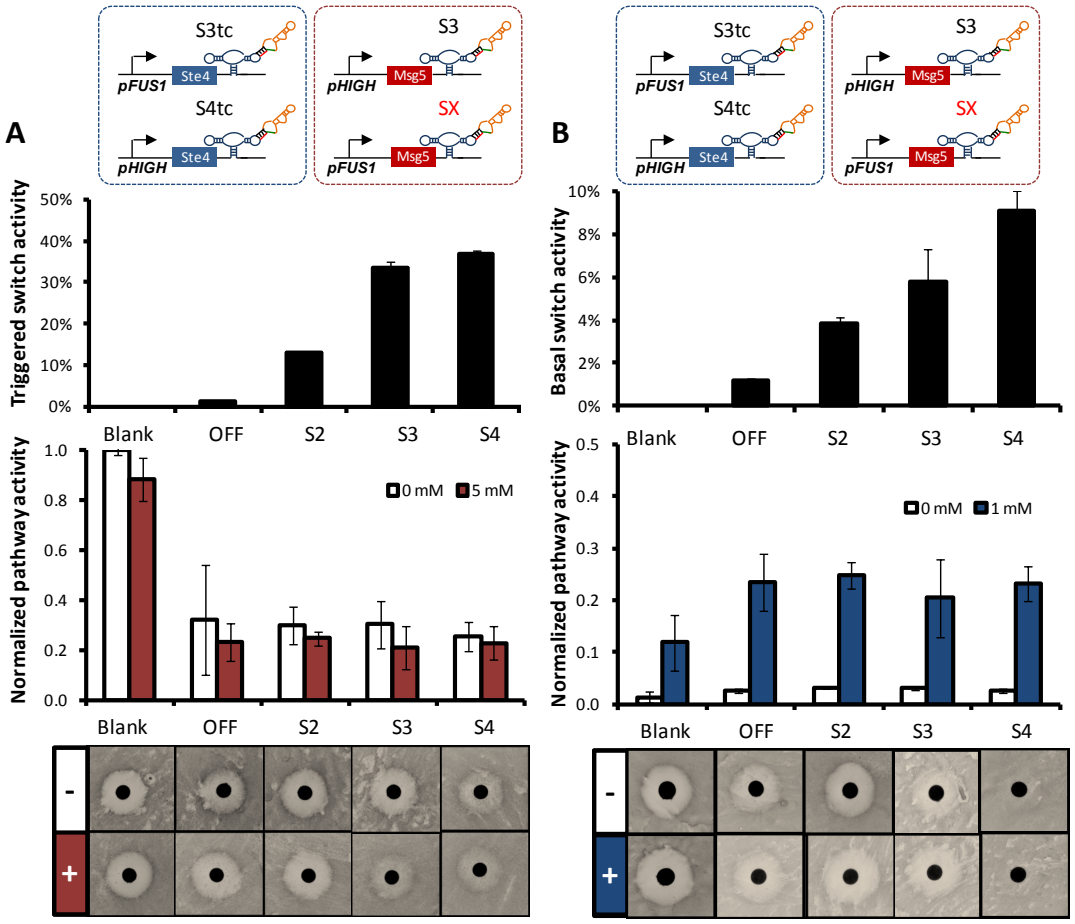


Figure 3.7. Networks configured with an amplifying diverter and various attenuating diverters show that pathway attenuation is a weak function of the strength of the negative feedback module. A. An amplifying diverter composed of pFUS1-Ste4-S3tc and pHIGH-Ste4-S4tc is paired with attenuating diverters incorporating pHIGH-Msg5-S3 and various strength negative feedback modules. While only modest changes are observed across the range of strengths for the negative feedback module, the attenuating diverter effectively routes to the chaste fate in the presence of the theophylline input when the active feedback strength of the negative diverter is at or above S3 levels. "SC" denotes strong chaste routing in the plate assays. **B.** The positive diverter weakly routes to the promiscuous fate in the presence of the tetracycline input when the basal feedback strength of the negative diverter is below S3. The basal feedback strength of S3 in the negative diverter is sufficient to inhibit the positive diverter.

The incorporation of an attenuating diverter, pairing both a resistance module and a negative feedback module, allowed chaste routing in the presence of the amplifying diverter. We next evaluated the ability of this dual diverter architecture to activate the pathway and route cells to the promiscuous fate in the presence of the positive diverter trigger over a range of negative feedback strengths. While GFP levels indicate that the amplifying diverter achieves moderate pathway activation within this network configuration, routing of fate to the promiscuous phenotype is weak across the range of negative feedback modules (Figure 3.7B). Based on these results, we chose to utilize a negative feedback module in the dual diverter architecture that provided strong chaste routing (pFUS1-Msg5-S3) as we examined tuning other components in the network to enhance pathway activation and routing to the promiscuous fate.

The performance of the amplifying and attenuating diverters in dual diverter networks are sensitive to the strength of the resistance module

To enhance routing to the promiscuous fate in the dual diverter network, we examined the impact of tuning the strength of the resistance module. The amplifying diverter (positive feedback: pFUS1-Ste4-S3tc; booster: pHIGH-Ste4-S4tc) was paired with attenuating diverters composed of a negative feedback module (pFUS1-Msg5-S3) and resistance modules incorporating RNA switches exhibiting a range of activities (pHIGH-Msg5-Sx). The performance and fate routing function of the dual diverter networks with varying resistance modules were characterized. For this dual diverter configuration, the amplifying diverter significantly increases pathway activity when

paired with weaker resistance modules (S2 and below). Pathway activation is reduced and the amplifying diverter is unable to strongly route cells to the promiscuous fate when paired with resistance modules of greater activity (S3 and above) (Figure 3.8). The results demonstrate that to preserve chaste routing the basal resistance activity must be between S2 and S3, indicating a threshold of resistance within this configuration that permits promiscuous routing via the amplifying diverter.

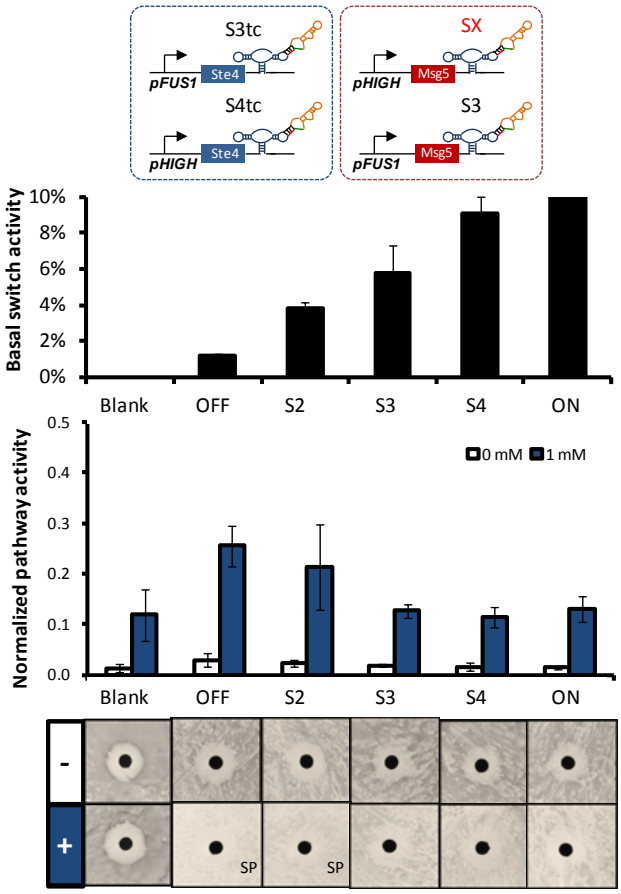


Figure 3.8. A dual-module positive diverter shows that pathway activation is sensitive to the activity of the resistance module. Amplifying diverter (positive feedback: pFUS1-Ste4-S3tc; booster: pHIGH-Ste4-S4tc) paired with various attenuating diverters (negative feedback: pFUS1-Msg5-S3; resistance: pHIGH-Msg5-Sx). "SP" denotes strong promiscuous routing in the plate assays.

We next evaluated the performance and fate routing of the attenuating diverter within the dual diverter network over varying resistance activities. The data demonstrate that pathway attenuation is sensitive to the activity of the resistance module (Figure 3.9).

Pathway activity drops as resistance increases from low to medium strength, above which the effect of resistance appears to saturate. Specifically, increasing the strength of resistance above that exhibited by the S3 module does not significantly reduce pathway activity. Strong chaste routing is observed in the presence of theophylline for resistance modules with activity at and above S3 levels. While no network configuration achieves dual-fate routing, the dual diverter network with S2 regulating the resistance module permits promiscuous routing while S3 enables chaste routing. These data indicate that dual-fate routing may be achieved by tuning the strength of the resistance module around the activation and attenuation thresholds that reside between S2 and S3.

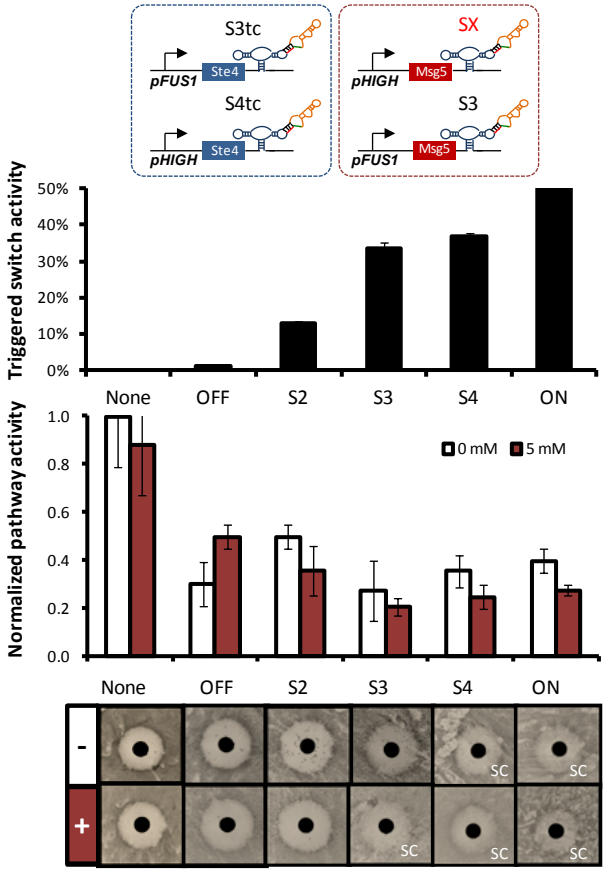


Figure 3.9. A dual-module negative diverter shows that pathway attenuation is a strong function of the activity of the resistance module. Amplifying diverter (positive feedback: pFUS1-Ste4-S3tc; booster: pHIGH-Ste4-S4tc) paired with various attenuating diverters (negative feedback: pFUS1-Msg5-S3; resistance: pHIGH-Msg5-Sx). "SC" denotes strong chaste routing in the plate assays.

The performance of attenuating and amplifying diverters is sensitive to the strength of the positive feedback module

We also examined the impact of tuning the strength of the positive feedback module on the dual-routing capability of the dual diverter network configuration. The attenuating diverter (resistance: pHIGH-Msg5-S3; negative feedback: pFUS1-Msg5-S3) was paired with amplifying diverters composed of a booster module (pHIGH-Ste4-S4tc) and positive feedback modules incorporating sets of RNA switches exhibiting varying activities (pFUS1-Ste4-Sx). We evaluated the ability of the attenuating diverter to reduce pathway activity and route cells to the chaste fate in this network configuration. Pathway attenuation increases modestly with decreasing positive feedback strength from S4tc to S2tc (Figure 3.10A). Across the range of positive feedback modules incorporating S2tc to S4tc, wild-type halo formation is mostly maintained in the absence of theophylline. For the network incorporating the OFF state control within the positive feedback module, pathway activity is low and cells are strongly routed to the chaste fate even in the absence of theophylline. While decreasing the strength of the positive feedback module may improve chaste routing, the data indicate that within this network configuration positive feedback above a particular threshold is required to maintain halo formation in the absence of either small-molecule trigger.

We next evaluated the ability of the amplifying diverter to increase pathway activation in these networks over the range of positive feedback strengths. Pathway activation increases with increasing positive feedback strength from OFF to S2tc. Increasing positive feedback strength above S2tc did not result in significant changes to either pathway activation or promiscuous routing within this network configuration

(Figure 3.10B). However, the wild-type halo was diminished in the absence of tetracycline over the range of feedback strengths compared to the same assay conditions (Figure 3.10A and B). These data indicate that the dual diverters' ability to maintain the wild-type halo may be sensitive to small, unaccounted for variations in plating conditions such as humidity and temperature which may affect plate diffusivity and thus the pheromone gradient. We additionally investigated whether reducing the positive feedback strength from S3tc to S2tc enhanced dual-fate routing in the dual diverter network. Specifically, we paired the amplifying diverter (booster: pHIGH-Ste4-S4tc; positive feedback: pFUS1-Ste4-S2tc) with attenuating diverters composed of a negative feedback module (pFUS1-Msg5-S3) and resistance modules incorporating a set of RNA switches exhibiting a range of activities. While reducing the strength of the positive feedback module did enhance chaste routing for the dual diverter network, it also resulted in weaker promiscuous routing even when configured with low activity resistance modules (Supplementary Figure 3.3). Reducing the strength of the booster module (pHIGH-Ste4-S3tc) similarly resulted in weaker promiscuous routing in this network configuration without significantly improving routing to the chaste fate (Supplementary Figure 3.4). The results indicate that the strength of the positive feedback module represents a sensitive parameter in the dual diverter network that must be precisely tuned to allow robust function of the dual-routing activities.

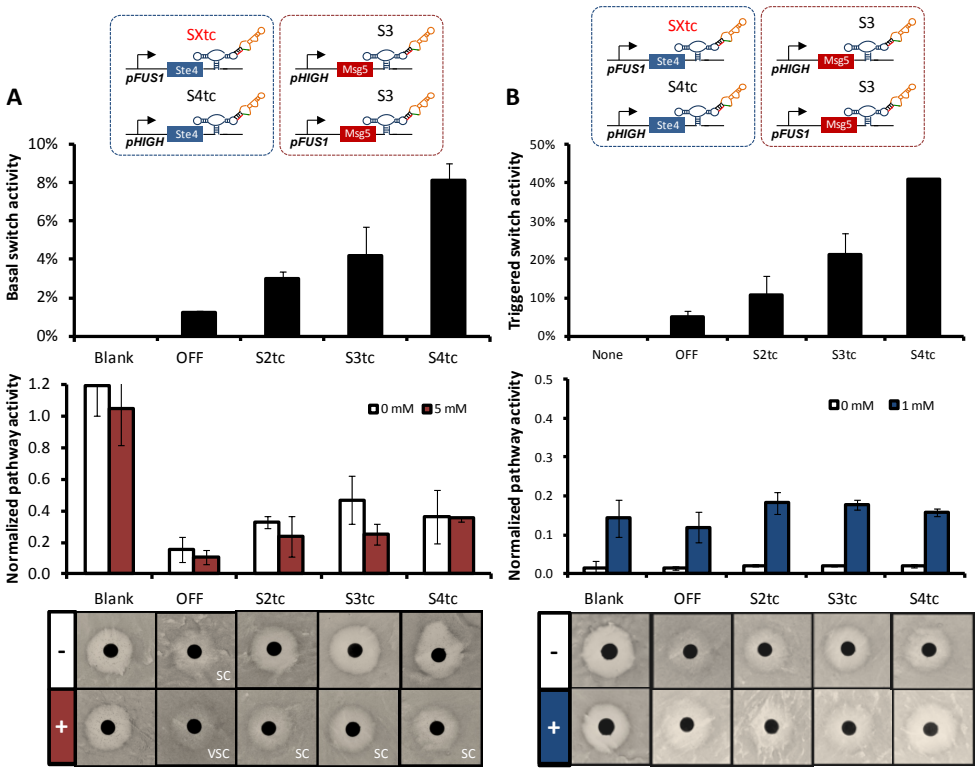


Figure 3.10. A dual-module negative diverter shows that pathway attenuation and activation are sensitive to the strength of the positive feedback module. A. Attenuating diverters (negative feedback: pFUS1-Msg5-S3; resistance: pHIGH-Msg5-S3) paired with various amplifying diverters (positive feedback: pFUS1-Ste4-SXtc; booster: pHIGH-Ste4-S4tc). "SC" and "VSC" denotes strong and very strong chaste routing in the plate assays, respectively. **B.** Pathway activation is relatively insensitive to varying positive feedback strength and only very weak routing is observed across the range of networks. However, insufficient levels of positive feedback in this network configuration inhibit normal halo formation in the absence of either trigger.

Benchmarking network diverter performance

Having examined a variety of network configurations, we identified several configurations that achieve near-optimal performance given our set of parts and the architectures we had constructed. We sought to benchmark four of the best dual-routing configurations to networks with reduced diverter antagonism. To benchmark performance of the dual diverter networks, we compared the routing efficiency and efficacy of the dual diverters to minimally antagonistic alternative architectures (MAAAs). MAAAs

represent best-case scenarios for routing under conditions of reduced interference from the opposing diverter. Six different MAAAs, three positive and three negative, were composed to evaluate the performance of both the positive and negative diverters in various configurations.

Positive MAAAs provide a tool for benchmarking a dual diverter's efficiency in activating the pathway and efficacy in routing to the promiscuous fate. Positive MAAAs are constructed by reducing the strength of modules within the negative diverter, thereby minimizing the antagonistic impact exerted by the negative diverter on the function of the positive diverter. We compared the constructed dual diverters to three different positive MAAAs: the low-resistance, the no-negative-feedback, and the low-resistance-no-negative-feedback MAAAs (For details of configuration see legend in Figure 3.11). We selected four dual diverter configurations that vary in the strength of positive feedback and resistance modules (positive feedback: pFUS1-pSte4-Sxtc; booster: pHIGH-Ste4-S4tc; negative feedback: pFUS1-Msg5-S3; resistance: pHIGH-Msg5-Sx). For simplicity, we refer to these dual diverters as: Diverter A, B, C, and D. Diverters A and B are composed with S3tc regulating the positive feedback module and the resistance module is regulated via S2 and S3, respectively. Diverters C and D mirror Diverters A and B, respectively, except S2tc is incorporated into the positive feedback module. These dual diverters are expected to decrease in positive routing and increase in negative routing performance from A to D.

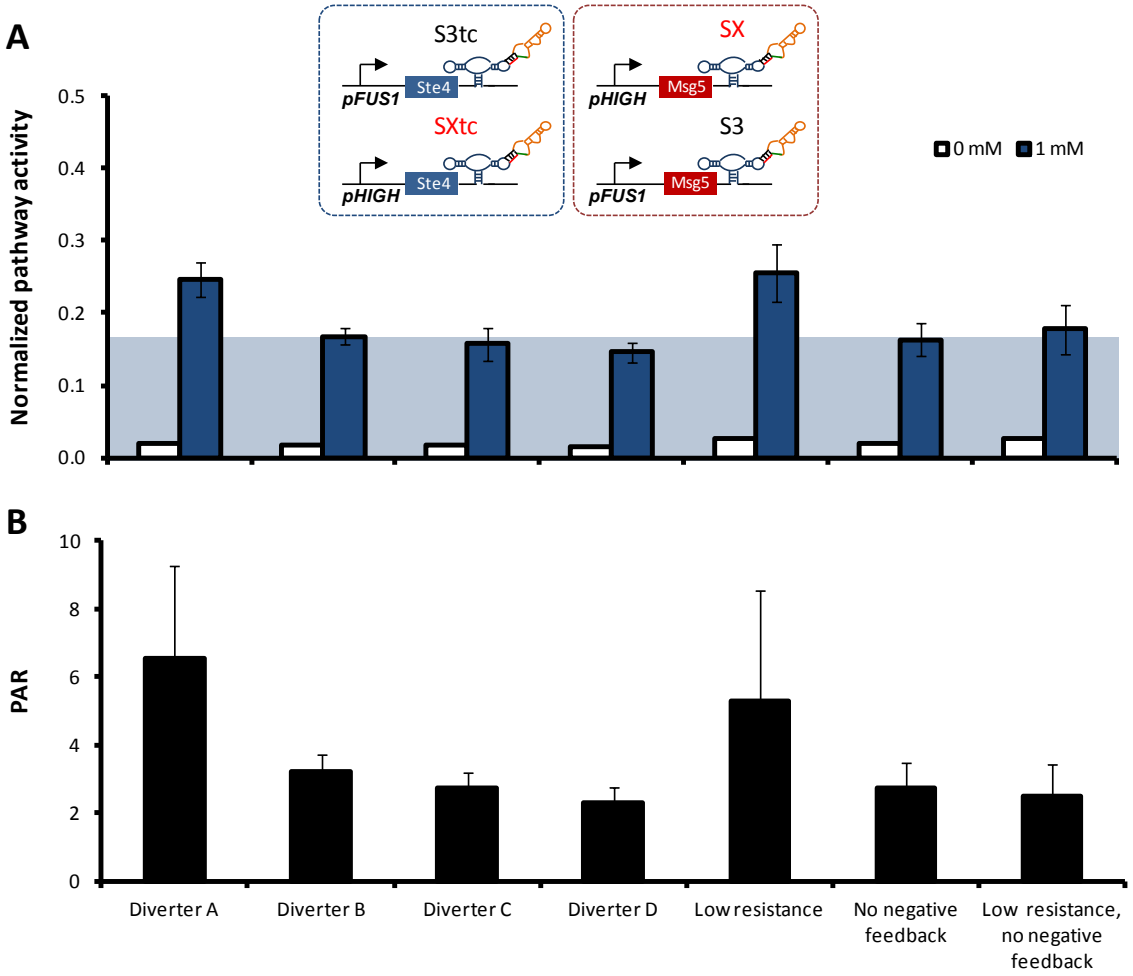


Figure 3.11. Benchmarking a dual diverters against various positive MAAAs indicates strong performance by Diverter A. A. Diverter A performs as well as the low resistance MAAA in activating the pathway, while pathway activation is notably diminished for Diverters B through D. All four dual diverter networks utilize a negative feedback module regulated by S3 and a booster module regulated via S4tc. Diverter A and Diverter B are composed with S3tc regulating the positive feedback module and the resistance module is regulated via S2 and S3, respectively. Diverters C and Diverter D mirror Diverters A and B respectively with S2tc substituted for S3tc in the positive feedback module. All the MAAAs containing the booster and negative feedback also have S4tc for the booster and S3 for negative feedback. Otherwise the MAAAs were composed without these elements to minimize antagonization. **B.** PAR values indicate that Diverter A performs as well as the MAAAs in differentiating the triggered and non-triggered populations, while performance drops for Diverter B and diminishes further across the range to Diverter D. The "low resistance" MAAA is composed with S3tc regulating positive feedback module, S4tc regulating the booster module, S3 regulating the negative feedback module and OFF regulating the resistance module. The low resistance MAAA represents the lowest possible level of resistance from this expression module given the other parameters (eg. strength of promoter, pathway regulator). The "no negative feedback" MAAA is composed similarly to the low resistance MAAA except it lacks a negative feedback module and contains S3 regulating the resistance module. The "no negative feedback, low resistance" MAAA lacks a negative feedback module and contains the OFF state control regulating the resistance module.

In evaluating positive routing, Diverter A performs as well as the low resistance MAAA in activating the pathway, while pathway activation is notably diminished for Diverters B through D (Figure 3.11A). As previously observed, pathway activation is particularly sensitive to the strength of resistance in the range between S2 and S3 when the positive feedback module is regulated by S3tc. Comparison of Diverters A and B with C and D reveals that decreasing the strength of the positive feedback module reduces this sensitivity by making the network less responsive to pathway activation. Interestingly, the positive MAAAs lacking a feedback module showed reduced pathway activity relative to the low-resistance MAAA. While lower pathway activity values for the MAAAs lacking a negative feedback module is unexpected, the differences in activity evaluated at a single time point may be the result of altered systems dynamics introduced by negative feedback. Negative autoregulation has been previously demonstrated to accelerate pathway response [25]. By evaluating a single time point of pathway activity, our assay may not capture an equally predictive measure of cellular fate when comparing different architectures. PAR values indicate that Diverter A performs as well as the low resistance MAAA in differentiating the triggered and non-triggered populations, while the PAR value drops for Diverter B and diminishes further in Diverters C and D (Figure 3.11B). While the PAR value for Diverter B (PAR 3.1) is lower than that for A (PAR 6.6) and the low-resistance MAAA, Diverter B outperforms the booster diverter (PAR 1.4) and positive feedback diverter (PAR 1.9) in differentiating triggered and non-triggered populations (Figure 3.4). These data indicate that with further tuning the constructed networks may provide strong resolution between divergent cell fates.

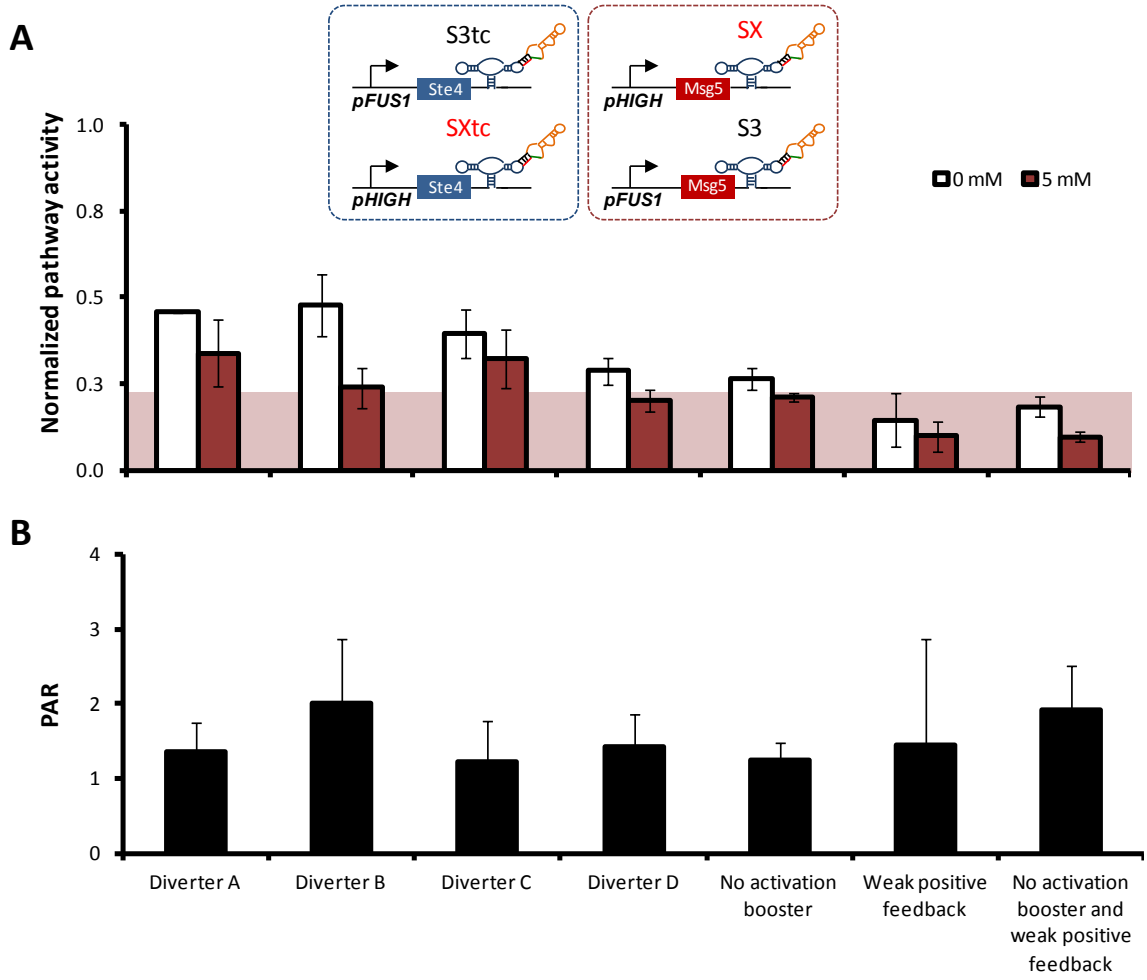


Figure 3.12. Benchmarking dual diverters against various negative MAAAs indicates strong performance from Diverter B and D. A. Normalized pathway activity data shows that pathway attenuation for Diverter B and D is similar to the no activation booster control. The data indicate that to further reduce pathway activity requires reducing the strength of the positive feedback module to the OFF state control levels. However, this configuration inhibits pathway activity even in the absence of theophylline. As observed previously, for these dual diverter networks a minimal level of positive feedback is necessary to maintain wild-type halo formation in the absence of either small molecule. **B.** PAR values indicate that Diverter B performs as well as or better than the negative MAAAs in separating triggered and non-triggered populations. The negative MAAAs were composed with S3tc regulating the positive feedback module, S4tc regulating the booster module, S3 regulating both the negative feedback and resistance modules with the following changes to the positive diverter. In the "no-activation-booster" MAAA lacks a booster module. The "weak-positive-feedback" MAAA is composed with the OFF state control regulating the positive feedback module. The no-activation-booster-weak-positive-feedback" MAAA lacks a booster module and contains a positive feedback module regulated by the OFF state control.

To benchmark the dual diverter's efficiency in pathway attenuation and efficacy of routing to the chaste fate, we constructed three negative MAAAs by reducing the strength of the expression modules in the positive diverter. We compared the four dual

diverters to the three negative MAAAs: no-activation-booster, weak-positive-feedback, and no-activation-booster-weak-positive-feedback MAAAs (For details of configuration see legend in Figure 3.12). When triggered with theophylline, pathway activation data shows that pathway attenuation for Diverters B and D is similar to the no activation booster control (Figure 3.12A). The data indicate that further reduction in pathway activity requires reducing the strength of the positive feedback module to levels exhibited by the OFF state control, which results in inhibition of pathway activity even in the absence of theophylline. As observed previously, a minimal level of positive feedback is necessary to maintain wild-type halo formation in the absence of either small-molecule trigger. From the measured PAR values, Diverter B performs as well as or better than the negative MAAAs in separating triggered and non-triggered populations (Figure 3.12B) The data indicate that configuring the dual diverter networks for differentiating between triggered and non-triggered cells and conditional chaste routing requires precise balancing of the positive feedback strength.

In comparing the various MAAAs and dual diverter configurations, Diverter B emerged as the best overall dual diverter. Diverter B achieved moderate levels of pathway activity and the second highest PAR value for pathway activation. Additionally, Diverter B attenuated pathway activity nearly as well as the no-booster MAAA and had a significantly higher PAR value for attenuation compared to the other dual diverters. While Diverter B represents the best overall dual diverter, Diverter A demonstrates that strong pathway activation can be achieved from the dual diverter architecture with amplifying and attenuating diverters simultaneously integrated. In addition, Diverter D demonstrates that strong pathway attenuation is possible from the same network

architecture tuned with different modular parts. Taken together, these results indicate that given the existing parts used in this study, we have optimized the dual routing system. Apart from modifying diverter configurations, further improvements in dual-fate routing may be achieved by increasing the trigger signals that activate the diverter switches and/or modifying growth conditions to modulate metabolic rate.

Conditional routing of genetically identical cells to diverse phenotypes in response to distinct molecular signals

We examined the performance of the dual diverters under conditions of elevated switch activation by examining dual-fate routing at elevated concentrations of the small-molecule triggers. Initial characterization of the dual diverter configurations were performed at extracellular concentrations of 5 mM theophylline and 1 mM tetracycline, which correspond to the standard concentrations for characterization of the RNA-based switches. However, increasing the concentration of small molecule input has been shown to increase the activation and thus level of expression from the RNA switches (Liang, J, et al. *Submitted*). To determine if increased activation of the theophylline- and tetracycline-responsive switches improved routing to the chaste and promiscuous fates, respectively, we performed halo assays in the absence and presence of slightly elevated concentrations of the environmental triggers (2 mM tetracycline, 20 mM theophylline). The results show that Diverters B, C, and D robustly route to the chaste fate when triggered with theophylline (Figure 3.13). Additionally, all four diverters modestly route cells to the promiscuous fate when triggered with tetracycline. However, Diverters B, C,

and D do not maintain wild-type halo formation in the absence of either trigger, instead demonstrating significant growth in the halo region in the absence of either trigger. We hypothesized that the positive diverter is overwhelmed by the negative diverter even at basal levels of expression resulting in reduced sensitivity to pheromone even in the absence of theophylline.

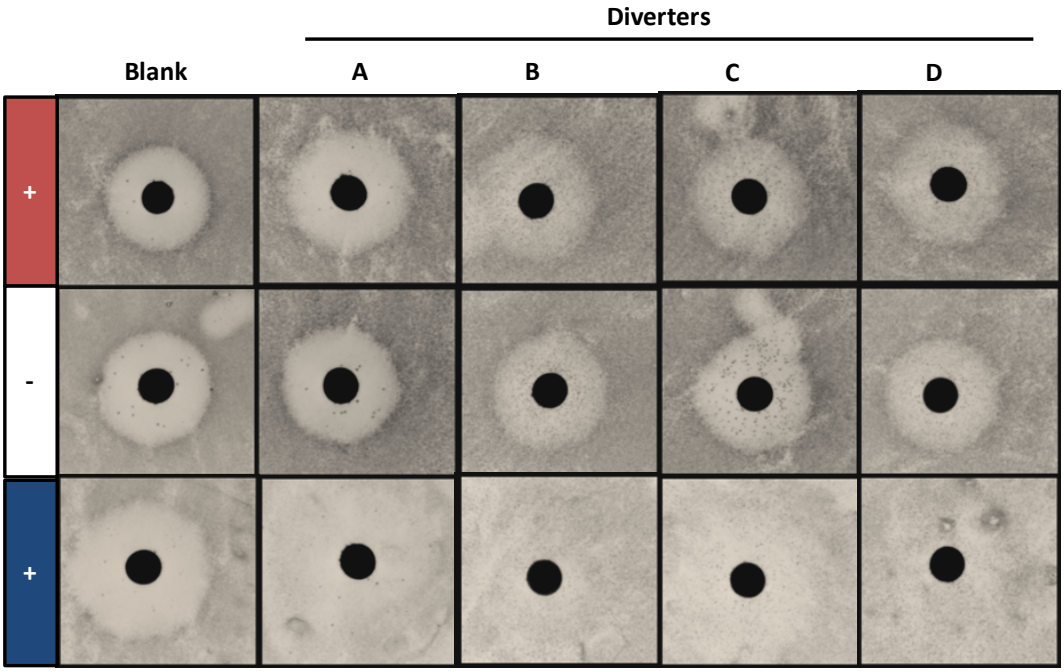


Figure 3.13. Higher small-molecule inputs improve dual-fate routing. Halo assays indicate that pushing the small-molecule concentration higher improves routing to both the promiscuous and chaste fate, but dual diverters routing to both fates (Diverters B, C, D) fail to maintain a robust wild-type halo in the absence of either trigger.

We attempted to restore the wild-type response in our dual diverters by altering growth conditions to support the positive diverter. Previous work demonstrated that the stability of positive feedback loops is sensitive to cellular metabolism [11]. In particular, reducing metabolism by exchange of sugar sources was shown to allow positive feedback loops to maintain memory of transient stimuli whereas under rapid growth conditions memory was lost. From this example, we postulated that noninducing, nonrepressing

(NINR) conditions, which are known to reduce metabolism compared to growth in dextrose, would support the positive feedback loop and restore the wild-type halo in the absence of either trigger. For all four diverters, the wild-type halo was restored in NINR conditions (Figure 3.14). Further, NINR conditions enhance promiscuous routing in the presence of tetracycline.

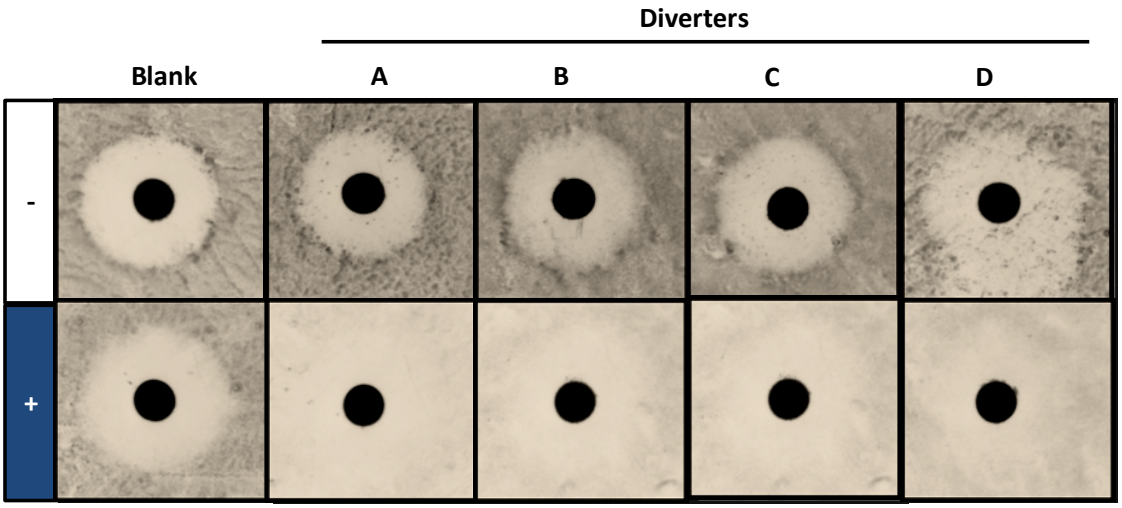


Figure 3.14. Metabolic modulation restores wild-type halo in the absence of either trigger and enhances promiscuous fate routing from dual-diverters. Halo assays performed in noninducing, nonrepressing (NINR) conditions demonstrate wild-type halos in the absence of either trigger. Assay perfomed at 0 mM (-) and 2 mM (+) tetracycline.

We further evaluated Diverter B, which emerged from the benchmarking as the best overall performing dual diverter, for dual routing at elevated small molecule levels supported by metabolic cues. As expected in NINR conditions, Diverter B maintains the wild-type halo until triggered by tetracycline to route cells to the promiscuous fate (Figure 3.15). When triggered with theophylline in dextrose, Diverter B routes cells to the chaste fate. Taken together, these data demonstrate that a dual diverter system has been optimized to achieve robust dual-fate routing from genetically identical cells in response to small-molecule triggers when supported by the appropriate metabolic cues.

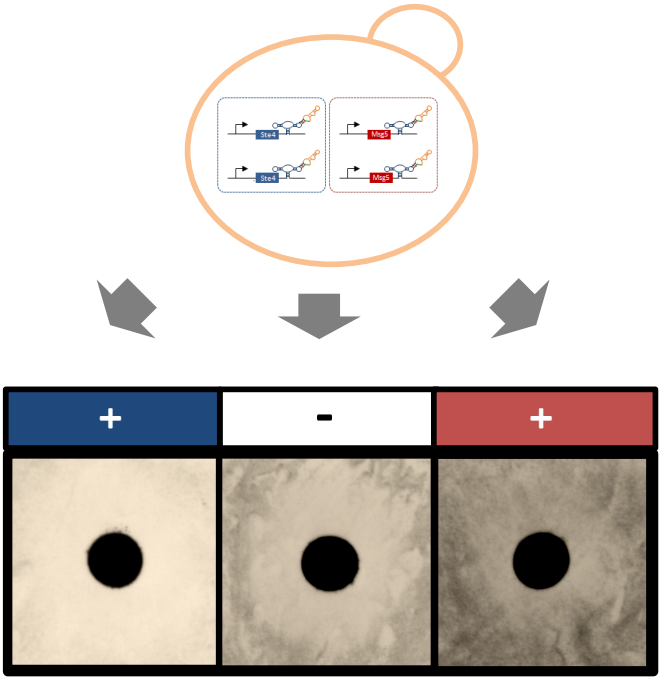


Figure 3.15. Routing genetically identical cells to divergent fates in response to small-molecule triggers supported by metabolic cues. Diverter B robustly routes cells to the promiscuous fate in response to tetracycline (left), preserved wild-type halo in absence of either trigger (center), and routes to the chaste fate in response to theophylline (right). Routing to the chaste fate is supported by dextrose, while wild-type and promiscuous fates are supported by NINR conditions.

Discussion

We have shown that molecular network diverters are capable of routing genetically identical cells to three divergent fates in response to specific environmental triggers when supported by metabolic modulation. The initial network configurations integrated single-module molecular network diverters, which had been previously optimized to route cell fate in the absence of the opposing diverter. Integration of the positive feedback diverter with the resistance diverter failed to permit dual-fate routing, and the basal expression levels from the opposing diverter resulted in antagonization, ultimately limiting the routing ability of each diverter. Addition of a booster module to the positive feedback diverter resulted in an amplifying diverter, which restored

promiscuous routing under low resistance conditions. In particular, the amplifying diverter paired with a low activity resistance module resulted in enhanced resolution between the triggered and non-triggered populations compared to the positive feedback diverter. Balancing the dual diverter networks with a negative feedback module facilitated routing to the chaste fate for some configurations at the sacrifice of promiscuous routing. By examining the network sensitivity to various parameters, our studies revealed that the strength of the positive feedback module and the activity of the resistance module represent highly sensitive control points. Our studies highlight the importance of architectures that allow for precise tuning of the diverter modules to allow dual fate routing. At lower levels of the environmental triggers none of the dual diverter networks achieved robust routing to both the promiscuous and chaste fates. However, under higher concentrations of the input triggers supported by metabolic modulation, a dual diverter configuration showed strong routing to both alternative fates while preserving the wild-type behavior in the absence of either trigger.

While we have optimized the dual diverter networks within the limits of the existing set of RNA switches, there remain potential opportunities for improving dual-fate routing performance. Given the sensitivity of the composed networks to the strength of the positive feedback module and the resistance module, balancing the strength of these modules represents an important design point for constructing networks that effectively route to both alternative fates. Basal level reduction from these expression modules represents a potential opportunity for further tuning of these networks.

Reducing expression from the resistance module and positive feedback module via promoter exchange or by modifying the expression context by integration or vector

exchange will reduce diverter antagonism from the opposing diverter by reducing basal expression levels. We used the pathway activity data for the dual diverter networks and switch expression levels to model the effect of reducing expression of both modules by one third for the Diverter B configuration (Figure 3.14). The results predict that absolute pathway activation and attenuation as well as PAR values for positive routing are improved in the modeled system relative to the best overall dual diverter, Diverter B. Thus, reducing expression from these cassettes may represent a potential opportunity for optimizing the performance of this dual diverter network configuration. In addition, integration at various loci, including GAL2 and TRP1, has been shown to reduce expression by 30–35%, providing a potential mechanism for reducing expression and variability from these constructs (Supplementary Figure 3.5). Promoter exchange represents another option for reducing expression from these modules. While there exist a variety of constitutive promoters with strengths lower than pHIGH, replacing a specific feedback promoter for variants of different strength may represent more of a challenge. However, previous work was performed to engineer variants of the mating promoter used in our study, pFUS1, that exhibit varying strengths of expression [26]. Several of the promoters demonstrated reduced strength and may be potential candidates for replacing the wild-type sequence used in these studies to reduce basal expression levels and enhance routing. While strategies that utilize promoter exchange or different expression contexts to tune basal levels will also result in reductions in the triggered levels of expression, addition of trans-acting controller may reduce basal expression without significantly reducing triggered expression.

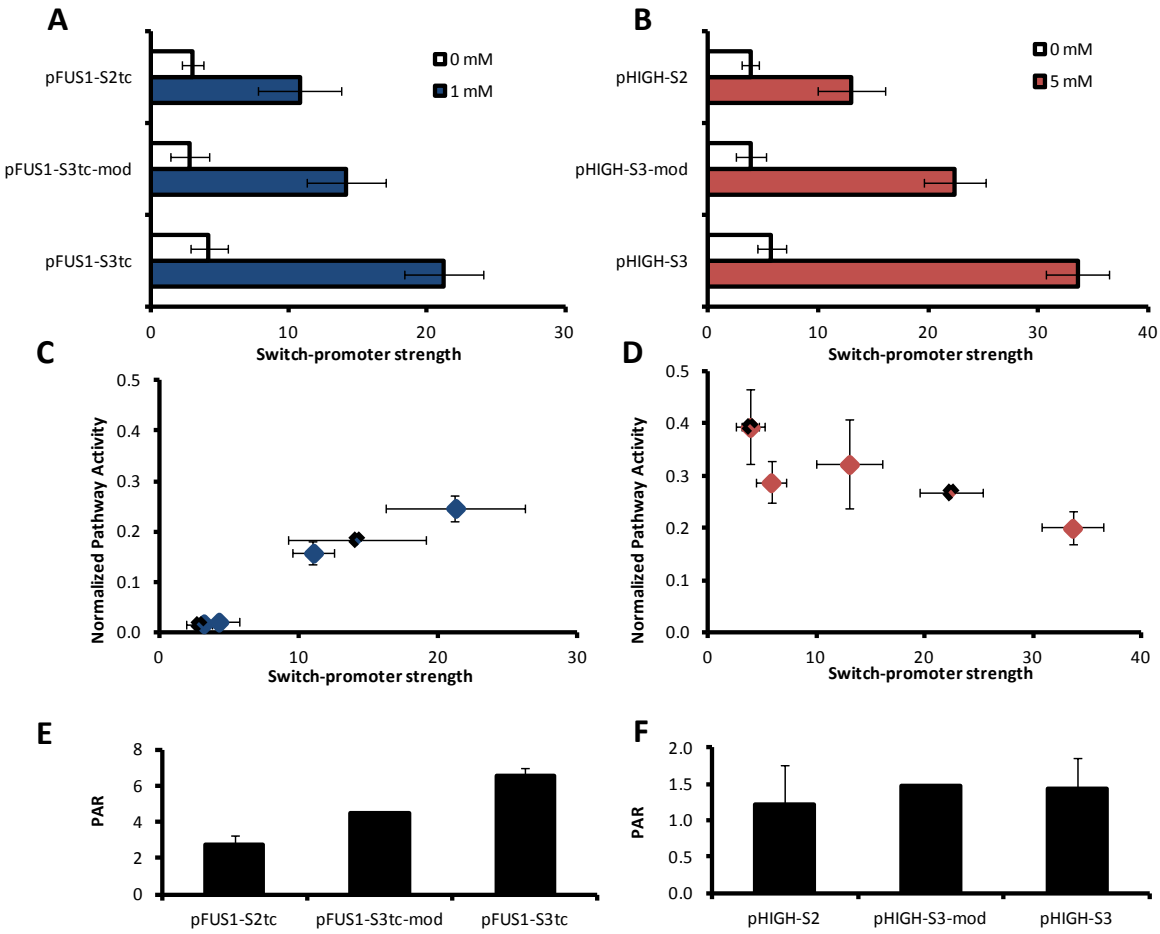


Figure 3.16. Reducing expression from the positive feedback and resistance modules in Diverter B may optimize the dual diverter network. A. Reducing expression from the feedback module by one third shifts the basal promoter-switch strength from S3tc levels to S2tc levels. pFUS1-S2tc denotes the feedback promoter combined with S2tc, pFUS1-S3tc denotes the feedback promoter with S3tc, and pFUS1-S3tc-mod indicates the promoter-switch strength for pFUS1-S3tc calculated to be reduced by one third. **B.** Reducing the resistance module by one third shifts basal levels from S3 to S2 levels. pHIGH-S2tc denotes the pHIGH promoter combined with S2tc, pHIGH-S3tc denotes pHIGH promoter with S3tc, and pHIGH-S3tc-mod indicates the promoter-switch strength for pFUS1-S3tc calculated to be reduced by one third. **C.** Pathway activity data are plotted in blue diamonds for pFUS1-S3tc (Diverter A) and pFUS1-S2tc (Diverter C) which both contain a resistance module regulated by S2. Diverter B contains a positive feedback module regulated via S3tc, a resistance module regulated via S3, a booster regulated via S4tc, and a negative feedback module. Pathway activity for the modified Diverter B was estimated by linearly interpolating between the values of pathway activity for pFUS1-S3tc (Diverter A) and pFUS1-S2tc (Diverter C) values of pathway activity, diamonds. **D.** Pathway activity data are plotted in red diamonds pHIGH-S2 (Diverter C) and pHIGH-S3 (Diverter D); both diverters contain S2tc regulating the positive feedback module. Pathway activity for the modified Diverter B was estimated by linearly interpolating between the values of pathway activity for pHIGH-S2 (Diverter C) and pHIGH-S3 (Diverter D), black diamonds. **E.** PAR values for pathway activation were determined from the measured and estimated values of pathway activity for the various configurations as previously described. **F.** PAR values for pathway attenuation for the various configurations were calculated using from measured and estimated values of pathway activity.

Trans-acting regulators of expression may provide a means for reducing basal expression levels, while being structured for minimal impact on the levels of triggered expression. Trans-acting RNA-based regulators, such as microRNAs or trans-acting ribozymes, offer a secondary layer of control by which to reduce basal level antagonization. Layering of trans-acting RNAs has been demonstrated to control basal level leakage, improving the performance of synthetic circuits [5, 12, 13]. Imbedding trans-acting RNA-based regulators within the transcripts of opposing regulators may facilitate the construction of mutually inhibitory loops between the two diverters. Such mutually inhibitory loops may be expected to increase the degree of divergence between cell fates and enhance the robustness of routing. Additionally, trans-acting RNA-based switches have been demonstrated to conditionally regulate target expression in mammalian cells [27]. Potentially, such trans-acting regulatory tools could provide a mechanism to conditionally reduce basal level expression with minimal effect on the triggered levels of expression. In addition to routing cells to three divergent fates, our work has demonstrated design principles for structuring networks to differentiate cells based on environmental cues and enhance the robustness of integrated mutually antagonistic programs. The dual diverter configurations incorporating the amplifying and attenuating diverters are structured to enhance routing while minimizing impact on performance of the opposing diverter. Achieving this effect required constructing functionally redundant genes with differential regulatory regions. Differential regulation of functionally redundant parts is a strategy that has been observed in natural biological systems, such as in the networks regulating bone formation and osteoblast differentiation as well as plant defense mechanisms and metabolism [28, 29]. Differential regulation of

functionally redundant genes may represent a common motif for amplifying pathway response to environmental cues that mediate changes in particular cellular behaviors and fate. Within our synthetic systems, we also observed that addition of a resistance module to the positive feedback diverter enhanced the resolution between the two populations of cells by reducing basal expression levels. We postulate that the resistance module enabled greater separation of the populations by buffering subthreshold noise of Ste4 expression. Addition of the booster module increased the levels of pathway activity in the network while preserving population resolution, providing a parameter for tuning the network output in this system. Synthetic networks that are structured to suppress noise amplification while amplifying differences in exogenous input may improve the selection of new RNA-based controllers by enhancing the resolution between switching and nonswitching elements. Additionally, when applied to cellular decision-making pathways, these enhanced amplifying circuits may facilitate robust differentiation between divergent cell fates.

We have demonstrated a scalable, modular, and tunable method for constructing RNA-based control systems that interact with a native signaling pathway to direct cell fate that may be readily translated to new pathways. As the synthetic biology tool box continues to expand, an increasingly number of synthetic control systems will be connected to native pathways, enabling more sophisticated and complex control for a wide array of applications. Many of the goals tissue engineering and molecular medicine such as the *ex vivo* construction of immunologically compatible tissues and the development of cell-based therapies will be advanced by synthetic control systems that spatially and temporally program cell fate.

Materials and methods

Plasmid construction

Standard molecular biology cloning techniques were used to construct all plasmids [30]. DNA synthesis was performed by Integrated DNA Technologies (Coralville, IA). All enzymes, including restriction enzymes and ligases, were obtained through New England Biolabs (Ipswich, MA). Ligation products were electroporated with a GenePulser XCell (Bio-Rad, Hercules, CA) into an *E. coli* DH10B strain (Invitrogen, Carlsbad, CA), where cells harboring cloned plasmids were maintained in Luria-Bertani media containing 50 mg/ml ampicillin (EMD Chemicals). All cloned constructs were sequence verified by Laragen Inc (Santa Monica, CA).

To construct the dual expression module negative and positive molecular network diverters, expression modules were cloned into dual cassette plasmids pCS2094 (Liang, J, et al. *Submitted*). pCS2094 served as the expression plasmid for positive feedback expression modules and resistance expression modules (Supplementary Figure 3.6A, Supplementary Table 3.1). The positive feedback expression module (pFUS1-Ste4-S2tc) was amplified via PCR with primers pFUS1.ClaI.pCS2094.FWD (5'-CCAATCTCAGAGGCTGAGTCTC) and Switch3'.XhoI.pCS2094.REV (5'-AAAACTCGAGTTTTTATTTTTCTTTTTGCTGTTTCG) and cloned into the unique ClaI and XhoI sites in pCS2094 to construct pKG227 (Supplementary Figure 3.6B). Tetracycline-responsive switches and appropriate controls (Supplementary Table 3.5) were inserted into the 3' UTR via the unique restriction sites AvrII and XhoI, located immediately downstream of the Ste4 stop codon as described previously [31]. Resistance expression modules (pTEF7-Msg5-SX) were PCR amplified from previously constructed

single-module molecule network diverter plasmids (Supplementary Table 2.6) using pTEF7.FWD (5'-AAGAGCTCATAGCTTCAAAATGTCTCTACTCCTTTTT) and CYC1t.NotI.REV (5'-AAAAGCGGCCGCTATATTACCCTGTTATCC) and cloned into the unique restriction sites SacI and NotI of construct harboring the positive feedback expression modules (Supplementary Table 3.2). To construct a complimentary expression system to pCS2094 (-URA) for expression of negative feedback expression modules and booster expression modules a secondary –TRP plasmid with dual expression cassettes was constructed. pKG233, a pCS2094 based-plasmid, was digested with SacI and KpnI, the dual-cassette fragment was gel extracted, and inserted via these same unique restriction sites in pCS1128 to compose pKG243 (Supplementary Figure 3.6C). Negative feedback expression modules (pFUS1-Msg5-SX) were constructed by PCR amplifying Msg5 and theophylline switches from previously constructed single-module negative feedback diverter plasmids (Supplementary Table 2.7) using Msg5.K2.FWD (5'-AAAGGATCCAATTAATAGTGCACATGCAATTTCAC) and Switch3'.XhoI.pCS2094.REV and cloned via the unique sites BamHI and XhoI. Booster expression modules (pTEF7-Ste4-SXtc) were added to the plasmids bearing the negative feedback expression modules via PCR of previously constructed booster diverter plasmids with pTEF7.FWD and CYC1t.NotI.REV and cloned into the unique restriction sites SacI and NotI (Supplementary Table 3.2). Single-module booster expression plasmids were constructed from pCS1128 for pairing with pCS2094. Previously constructed booster diverter plasmids were SacI and KpnI digested, the expression cassette was gel extracted, and cloned via these same unique restriction sites in pCS1128 (Supplementary Table 3.3).

Measuring mating pathway activity via a transcriptional reporter

The molecular diverter plasmids and appropriate controls were transformed into the previously constructed yeast mating reporter strain CSY840 (Supplementary Figure 3.9). Cells were inoculated into the appropriate dropout media, grown overnight at 30°C, and back diluted into fresh media in the presence or absence of ligand at the specified concentration to an OD_{600} of <0.1. For negative diverters, after growing for 3 hr at 30°C, cells were stimulated with saturating pheromone levels, to a final concentration of 100 nM α mating factor acetate salt (Sigma-Aldrich, St. Louis, MO), to activate the mating pathway. Following 6 hr of growth post-back-dilution, GFP fluorescence levels from the pFUS1-yEGFP3 reporter were evaluated via flow cytometry using Cell Lab Quanta SC flow cytometer (Beckman Coulter, Fullerton, CA) with the following settings: 488-nm laser line, 525-nm bandpass filter, and photomultiplier tube setting of 5.0 on FL1 (GFP). Fluorescence data were collected under low flow rates for ~ 10,000 viable cells. Normalized pathway activity is calculated as the geometric mean of three biological replicates of each sample normalized to the blank plasmid control stimulated with saturating α mating factor in the absence of either small molecule.

Measuring mating pathway activity via halo assays

Mating associated cell-cycle arrest was evaluated via halo assays [32]. Halo assays were performed on cultures grown overnight in YNB with 2% dextrose or in 1% sucrose, 2% raffinose for NINR noninducing, nonrepressing (NINR) conditions and appropriate dropout solution. Overnight cultures were back diluted into fresh media and

grown to OD 600 ~ 0.2–0.4. 200 μl of each replicate were plated on the appropriate dropout plates containing no small molecule, theophylline or tetracycline at the specified concentration on dextrose except where specifically indicated for NINR conditions. Standard concentrations are defined as 5 mM for theophylline and 1mM for tetracycline. After plating the cells, a gradient of α mating factor was established by saturating a filter disk (2 mm diameter) of Whatman paper with 9 μl of 0.1 mg/mL α mating factor and placing the disk on the center of the plate. Cells were grown for 18–24 hr at 30°C and imaged via epi-white illumination with a GelDoc XR+ System (Bio-Rad).

Acknowledgements

Thanks to Joe Liang for plasmid pCS2094 as well as technical advice in incorporating and characterizing various switches into the molecular network diverters. Funding provided by the National Institutes of Health (R01GM086663).

References

1. Perez-Moreno M, Fuchs E: **Catenins: keeping cells from getting their signals crossed.** *Dev Cell* 2006, **11:**601-612.
2. King TJ, Lampe PD: **Temporal regulation of connexin phosphorylation in embryonic and adult tissues.** *Biochim Biophys Acta* 2005, **1719:**24-35.
3. Schiller M, Javelaud D, Mauviel A: **TGF-beta-induced SMAD signaling and gene regulation: consequences for extracellular matrix remodeling and wound healing.** *J Dermatol Sci* 2004, **35:**83-92.
4. Sanjuan MA, Milasta S, Green DR: **Toll-like receptor signaling in the lysosomal pathways.** *Immunol Rev* 2009, **227:**203-220.
5. Callura JM, Dwyer DJ, Isaacs FJ, Cantor CR, Collins JJ: **Tracking, tuning, and terminating microbial physiology using synthetic riboregulators.** *Proc Natl Acad Sci U S A* 2010, **107:**15898-15903.
6. Kobayashi H, Kaern M, Araki M, Chung K, Gardner TS, Cantor CR, Collins JJ: **Programmable cells: interfacing natural and engineered gene networks.** *Proc Natl Acad Sci U S A* 2004, **101:**8414-8419.
7. Purnick PE, Weiss R: **The second wave of synthetic biology: from modules to systems.** *Nat Rev Mol Cell Biol* 2009, **10:**410-422.
8. Win MN, Smolke CD: **A modular and extensible RNA-based gene-regulatory platform for engineering cellular function.** *Proc Natl Acad Sci U S A* 2007.
9. Gardner TS, Cantor CR, Collins JJ: **Construction of a genetic toggle switch in Escherichia coli.** *Nature* 2000, **403:**339-342.

10. Friedland AE, Lu TK, Wang X, Shi D, Church G, Collins JJ: **Synthetic gene networks that count.** *Science* 2009, **324:**1199-1202.

11. Ajo-Franklin CM, Drubin DA, Eskin JA, Gee EP, Landgraf D, Phillips I, Silver PA: **Rational design of memory in eukaryotic cells.** *Genes Dev* 2007, **21:**2271-2276.

12. Deans TL, Cantor CR, Collins JJ: **A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells.** *Cell* 2007, **130:**363-372.

13. Xie Z, Wroblewska L, Prochazka L, Weiss R, Benenson Y: **Multi-input RNAi-based logic circuit for identification of specific cancer cells.** *Science* 2011, **333:**1307-1311.

14. Ferrell JE, Xiong W: **Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible.** *Chaos* 2001, **11:**227-236.

15. Becskei A, Seraphin B, Serrano L: **Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion.** *Embo J* 2001, **20:**2528-2535.

16. Becskei A, Serrano L: **Engineering stability in gene networks by autoregulation.** *Nature* 2000, **405:**590-593.

17. Mangan S, Itzkovitz S, Zaslaver A, Alon U: **The incoherent feed-forward loop accelerates the response-time of the gal system of Escherichia coli.** *J Mol Biol* 2006, **356:**1073-1081.

18. Mangan S, Alon U: **Structure and function of the feed-forward loop network motif.** *Proc Natl Acad Sci U S A* 2003, **100:**11980-11985.

19. Bashor CJ, Helman NC, Yan S, Lim WA: **Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics.** *Science* 2008, **319:**1539-1543.

20. Ferrell JE, Jr.: **Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability.** *Curr Opin Cell Biol* 2002, **14:**140-148.

21. Ferrell JE, Jr., Pomerening JR, Kim SY, Trunnell NB, Xiong W, Huang CY, Machleder EM: **Simple, realistic models of complex biological processes: positive feedback and bistability in a cell fate switch and a cell cycle oscillator.** *FEBS Lett* 2009, **583:**3999-4005.

22. Xiong W, Ferrell JE, Jr.: **A positive-feedback-based bistable 'memory module' that governs a cell fate decision.** *Nature* 2003, **426:**460-465.

23. Espinosa-Soto C, Padilla-Longoria P, Alvarez-Buylla ER: **A gene regulatory network model for cell-fate determination during Arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles.** *Plant Cell* 2004, **16:**2923-2939.

24. Santos SD, Verveer PJ, Bastiaens PI: **Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate.** *Nat Cell Biol* 2007, **9:**324-330.

25. Rosenfeld N, Elowitz MB, Alon U: **Negative autoregulation speeds the response times of transcription networks.** *J Mol Biol* 2002, **323:**785-793.

26. Ingolia NT, Murray AW: **Positive-feedback loops as a flexible biological module.** *Curr Biol* 2007, **17:**668-677.

27. Beisel CL, Chen YY, Culler SJ, Hoff KG, Smolke CD: **Design of small molecule-responsive microRNAs based on structural requirements for Drosha processing.** *Nucleic Acids Res* 2011, **39:**2981-2994.

28. Banerjee C, Javed A, Choi JY, Green J, Rosen V, van Wijnen AJ, Stein JL, Lian JB, Stein GS: **Differential regulation of the two principal Runx2/Cbfa1 n-terminal isoforms in response to bone morphogenetic protein-2 during development of the osteoblast phenotype.** *Endocrinology* 2001, **142:**4026-4039.

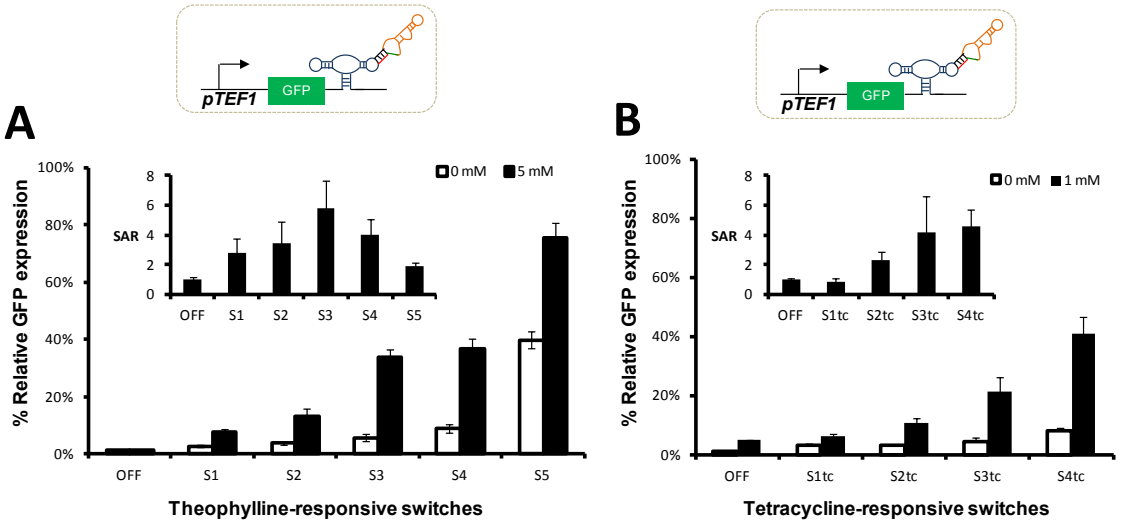
29. Yuan Y, Chung JD, Fu X, Johnson VE, Ranjan P, Booth SL, Harding SA, Tsai CJ: **Alternative splicing and gene duplication differentially shaped the regulation of isochorismate synthase in Populus and Arabidopsis.** *Proc Natl Acad Sci U S A* 2009, **106:**22020-22025.

30. Sambrook J RD: *Molecular Cloning: A Laboratory Manual, 3rd edn.* Cold Spring Harbor, NY: Cold Spring Harbor Lab Press; 2001.

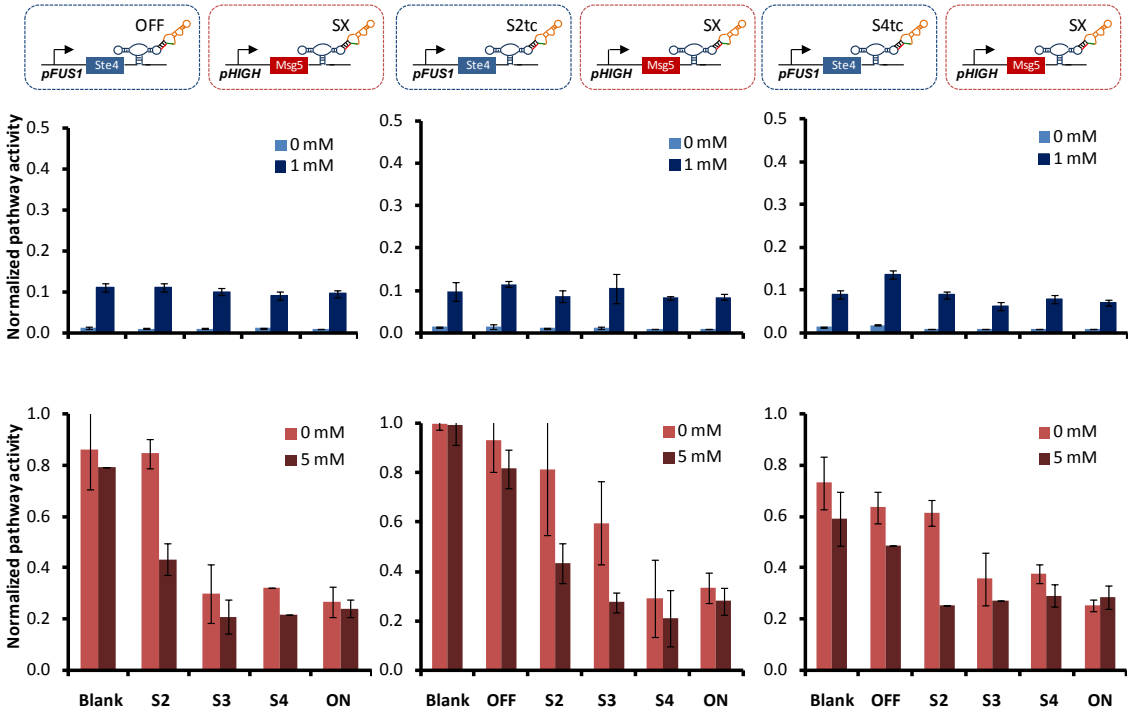
31. Win MN, Smolke CD: **A modular and extensible RNA-based gene-regulatory platform for engineering cellular function.** *Proceedings of the National Academy of Sciences* 2007, **104:**14283-14288.

32. Sprague GF, Jr.: **Assay of yeast mating reaction.** *Methods Enzymol* 1991, **194:**77-93.

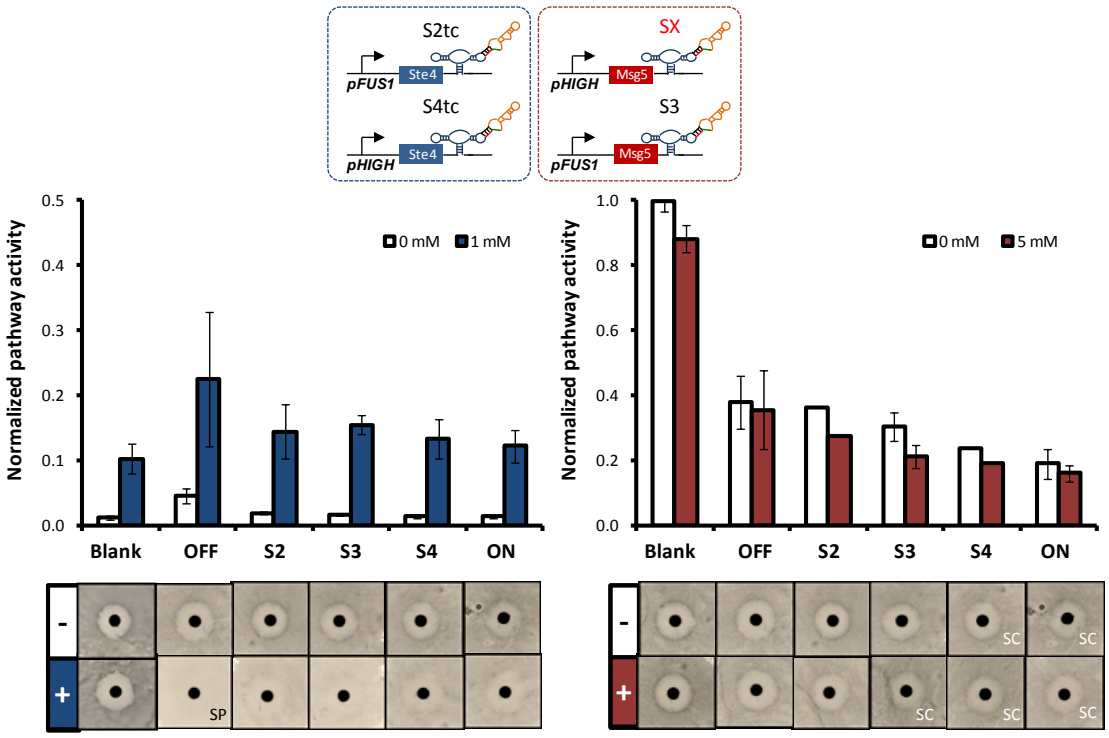
Supplementary figures



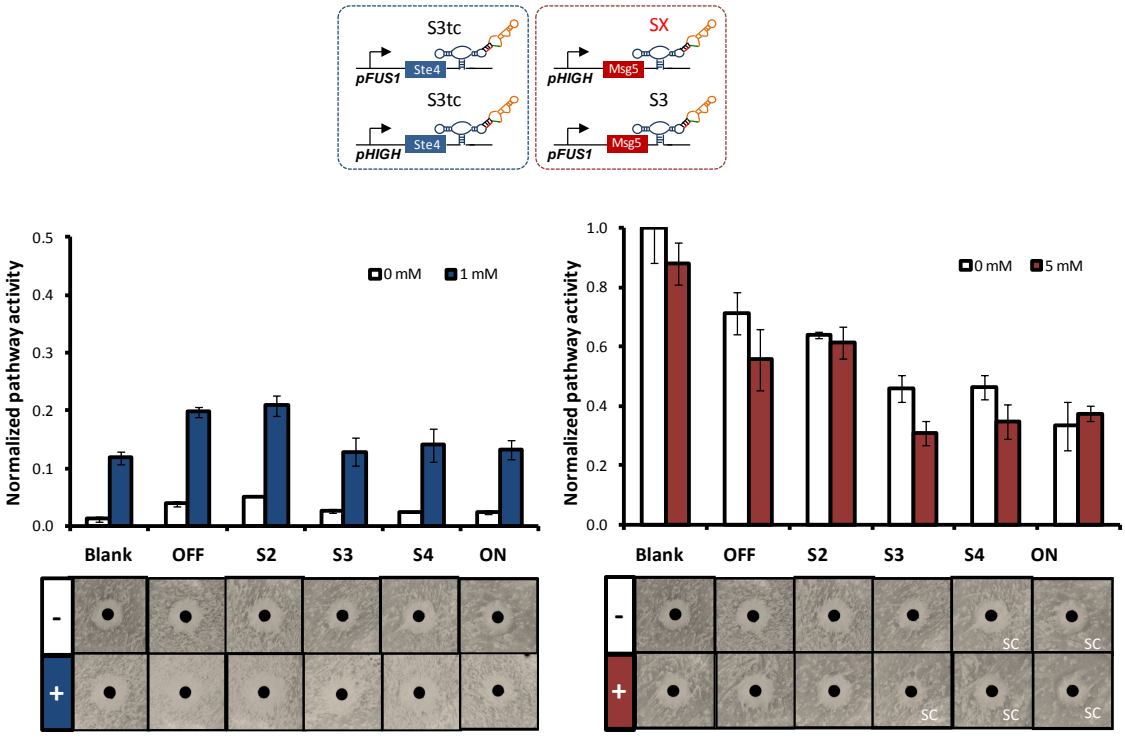
Supplementary Figure 3.1. Range of switch expression strengths. A. Theophylline-responsive switches' basal levels, the expression levels in the absence of ligand, range from 3% to 40% of the ON control. S3 has the highest switch activation ratio (SAR) at 5.7 . SAR is the ratio of expression levels in the presence of ligand to the level in the absence of ligand. **B.** Tetracycline-responsive switches' basal levels range from 3% to 8%. S3tc and S4tc have similar high SARs at 5.2 and 5.1, respectively.



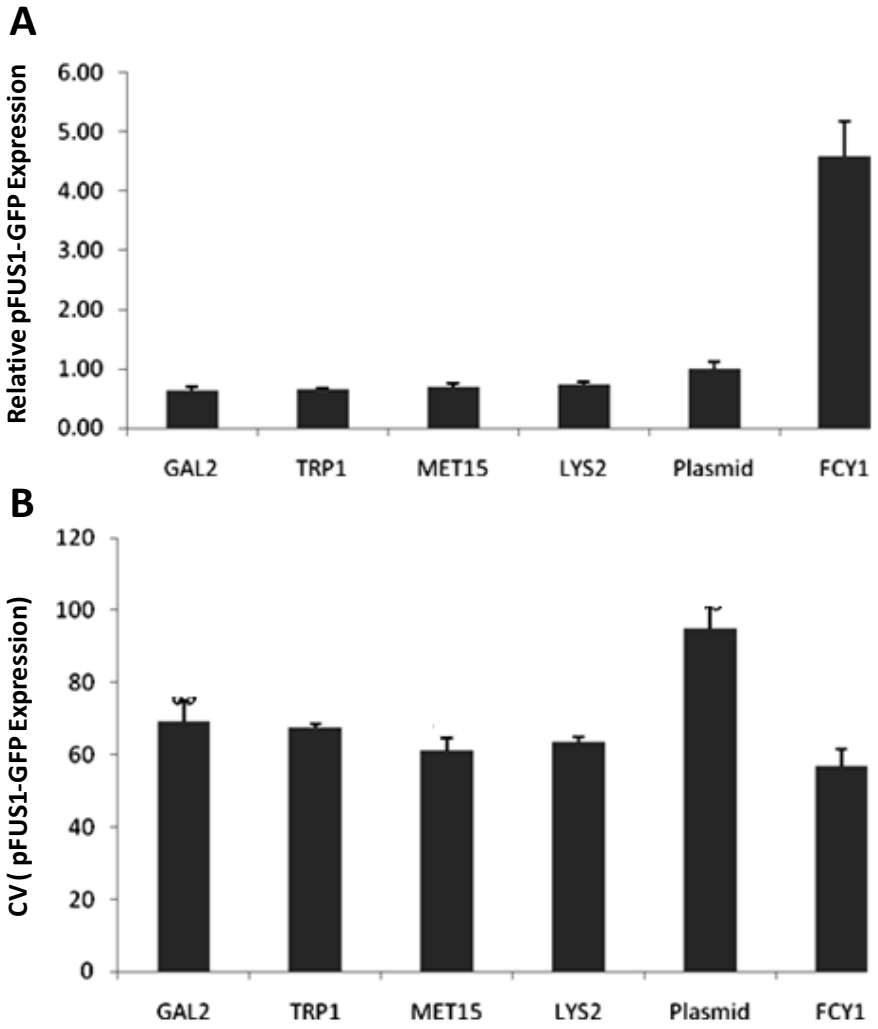
Supplementary Figure 3.2. Single-module diverters fail to achieve dual-fate routing. Single-module positive feedback diverters, pFUS1-Ste4 with OFF, S2tc, and S4tc, fail to route to the promiscuous fate in the presence of negative diverters incorporating a range of switch strengths. The Blank control bears a plasmid without either diverter.



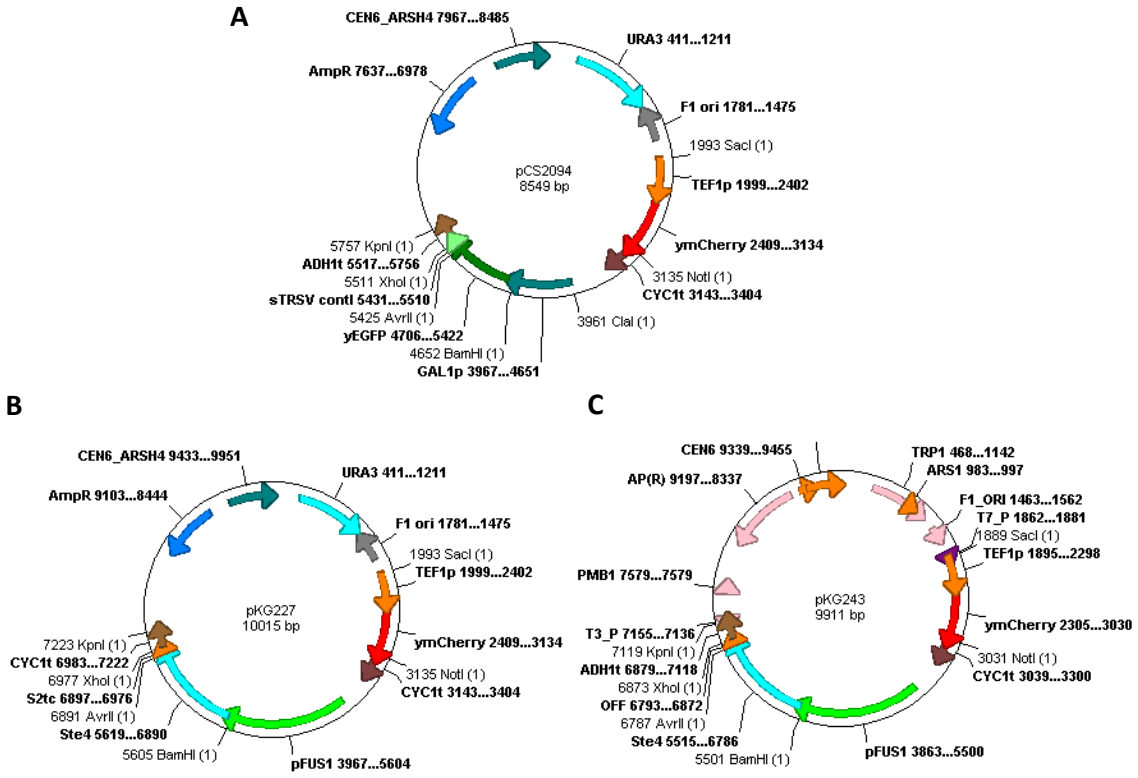
Supplementary Figure 3.3. Reducing the strength of the feedback module from S3tc to S2tc yields weak promiscuous routing while modestly improving chaste routing.



Supplementary Figure 3.4. Reducing the strength of the booster module from S4tc to S3tc yields weak promiscuous routing and does not significantly improve chaste routing.



Supplementary Figure 3.5. Loci characterization. A. Relative pFUS1-GFP expression indicates that integration generally reduces expression 25–35 % for the loci, except at the FCY1 locus which increases expression nearly 5-fold compared to the plasmid. The loci were characterized by flow cytometry of cells with pFUS1-GFP integrants stimulated with 100 nM for 3 hrs. **B.** Integration also reduced the coefficient of variation significantly.



Supplementary Figure 3.6. Plasmid maps: A. pCS2094 **B.** pKG227 **C.** pKG243

Supplementary tables

Supplementary Table 3.1. pCS2094-based dual-expression cassette plasmids

pCS #	pKG#	Downstream Cassette	Upstream Cassette	Marker
pCS2094	**205**	**pTEF1-yEGFP-ON-CYC1t**	**pTEF1-mcherry-CYC1t**	**URA**
	227	**pFUS1-Ste4-S2tc**	**pTEF1-mcherry-CYC1t**	**URA**
	306	pFUS1-Ste4-S2tc	pTEF7 Msg5 OFF	URA
	231	pFUS1-Ste4-S2tc	pTEF7 Msg5 S2	URA
	307	pFUS1-Ste4-S2tc	pTEF7 Msg5 S47	URA
	232	pFUS1-Ste4-S2tc	pTEF7 Msg5 S3	URA
	283	pFUS1-Ste4-S2tc	pTEF7 Msg5 ON	URA
	pKG233	**pFUS1-Ste4-OFF**	**pTEF1-mcherry-CYC1t**	**URA**
	pKG238	pFUS1-Ste4-OFF	pTEF7 Msg5 S2	URA
	pKG300	pFUS1-Ste4-OFF	pTEF7 Msg5 S47	URA
	pKG239	pFUS1-Ste4-OFF	pTEF7 Msg5 S3	URA
	pKG271	pFUS1-Ste4-OFF	pTEF7 Msg5 ON	URA
	pKG298	**pFUS1-Ste4-Stc-OFF1**	**pTEF1-mcherry-CYC1t**	**URA**
	pKG302	pFUS1-Ste4-Stc-OFF1	pTEF7 Msg5 OFF	URA
	pKG303	pFUS1-Ste4-Stc-OFF1	pTEF7 Msg5 S2	URA
	pKG304	pFUS1-Ste4-Stc-OFF1	pTEF7 Msg5 S47	URA
	pKG305	pFUS1-Ste4-Stc-OFF1	pTEF7 Msg5 S3	URA
	pKG316	pFUS1-Ste4-Stc-OFF1	pTEF7 Msg5 ON	URA
	pKG234	**pFUS1-Ste4-S3tc**	**pTEF1-mcherry-CYC1t**	**URA**
	pKG310	pFUS1-Ste4-S3tc	pTEF7 Msg5 OFF	URA
	pKG236	pFUS1-Ste4-S3tc	pTEF7 Msg5 S2	URA
	pKG313	pFUS1-Ste4-S3tc	pTEF7 Msg5 S47	URA
	pKG237	pFUS1-Ste4-S3tc	pTEF7 Msg5 S3	URA
	pKG280	pFUS1-Ste4-S3tc	pTEF7 Msg5 ON	URA
	pKG228	**pFUS1-Ste4-ON**	**pTEF1-mcherry-CYC1t**	**URA**
	pKG311	pFUS1-Ste4-ON	pTEF7 Msg5 OFF	URA

Supplementary Table 3.2. pCS1128-based dual-expression cassette plasmids

pCS #	pKG#	Downstream Cassette	Upstream Cassette	Marker
pCS1128	**-**	**None**	**pGAL1-yEGFP-CYC1t**	**TRP**
	pKG243	**pFUS1-Ste4-OFF**	**pTEF1-mcherry-CYC1t**	**TRP**
	pKG251	**pFUS1-Msg5-OFF**	**pTEF1-mcherry-CYC1t**	**TRP**
	pKG265	pFUS1-Msg5-OFF	pTEF7 Ste4 S2tc	TRP
	pKG308	pFUS1-Msg5-OFF	pTEF7 Ste4 Stc-OFF1	TRP
	pKG317	pFUS1-Msg5-OFF	pTEF7 Ste4 S3tc	TRP
	pKG267	pFUS1-Msg5-OFF	pTEF7 Ste4 ON	TRP
	pKG248	**pFUS1-Msg5-S2**	**pTEF1-mcherry-CYC1t**	**TRP**
	pKG273	pFUS1-Msg5-S2	pTEF7 Ste4 OFF	TRP
	pKG258	pFUS1-Msg5-S2	pTEF7 Ste4 S2tc	TRP
	pKG299	pFUS1-Msg5-S2	pTEF7 Ste4 Stc-OFF1	TRP
	pKG318	pFUS1-Msg5-S2	pTEF7 Ste4 S3tc	TRP
	pKG260	pFUS1-Msg5-S2	pTEF7 Ste4 ON	TRP
	pKG287	**pFUS1-Msg5-S47**	**pTEF1-mcherry-CYC1t**	**TRP**
	pkG288	pFUS1-Msg5-S47	pTEF7 Ste4 OFF	TRP
	pKG301	pFUS1-Msg5-S47	pTEF7 Ste4 S2tc	TRP
	pKG309	pFUS1-Msg5-S47	pTEF7 Ste4 Stc-OFF1	TRP
	pKG319	pFUS1-Msg5-S47	pTEF7 Ste4 S3tc	TRP
	pKG290	pFUS1-Msg5-S47	pTEF7 Ste4 ON	TRP
	pKG249	**pFUS1-Msg5-S3**	**pTEF1-mcherry-CYC1t**	**TRP**
	pKG291	pFUS1-Msg5-S3	pTEF7 Ste4 OFF	TRP
	pKG268	pFUS1-Msg5-S3	pTEF7 Ste4 S2tc	TRP
	pKG314	pFUS1-Msg5-S3	pTEF7 Ste4 Stc-OFF1	TRP
	pKG320	pFUS1-Msg5-S3	pTEF7 Ste4 S3tc	TRP
	pKG267	pFUS1-Msg5-S3	pTEF7 Ste4 ON	TRP
	pKG250	**pFUS1-Msg5-ON**	**pTEF1-mcherry-CYC1t**	**TRP**
	pKG261	pFUS1-Msg5-ON	pTEF7 Ste4 OFF	TRP
	pKG262	pFUS1-Msg5-ON	pTEF7 Ste4 S2tc	TRP
	pKG315	pFUS1-Msg5-ON	pTEF7 Ste4 Stc-OFF1	TRP

Supplementary Table 3.3. pCS1128-based single-module booster plasmids

pTEF7-Ste4-CYC1t			
pCS #	**pKG#**	**Parent**	**3'UTR**
	pKG321	pCS1128	S4tc
	pKG295	pCS1128	S3tc
	pKG188	pCS1128	S2tc
	pKG182	pCS1128	S1tc
	pKG181	pCS1128	OFF

Chapter 4

Development of trans-ribozymes as actuators controlling gene expression

Abstract

The modular and tunable nature of RNA makes it ideal to perform both sensing and actuation functions within an synthetic control system [1]. RNA regulates gene expression through a variety of mechanisms at the transcriptional, post-transcriptional, and translational levels. Ribozymes are RNA molecules that regulate gene expression through post-transcriptional cleavage of target RNA. While their ability to cleave transcripts in a sequence-specific manner makes trans-ribozymes attractive candidates for targeting therapeutically relevant transcripts, applications have been limited by poor *in vivo* efficiency. However, through the recent elucidation of the design rules for *in vivo* catalytic activity, hammerhead ribozymes are now poised to be effective regulators of gene expression. In this work, we apply the rules for *in vivo* activity established for cis-ribozymes to the development of more effective trans-ribozymes. We demonstrate that precise engineering of the intramolecular reaction between the ribozyme and target transcript is required for efficient cleavage at physiological Mg^{2+} concentrations. To improve the correlation between our *in vitro* and *in vivo* assays of ribozyme cleavage efficiency, we employed a dual cis-hammerhead ribozyme cassette to excise our trans-ribozyme designs from the ribonucleoprotein (RNP) *in vivo*. Finally, we varied the levels of target sequence to examine how the concentration of target transcript, and consequently, the ratio of ribozyme to target affect target knockdown levels. These experiments demonstrated that knockdown efficiency has a biphasic relationship to target transcript levels. Our results suggest that one of the key limitations to trans-ribozyme-mediated knockdown may be facilitating the intermolecular binding event between the trans-acting ribozymes and the target transcript. While this limits the potential application

space of trans-acting ribozymes, it may position these devices as uniquely targeted regulators of endogenous transcripts associated with cell proliferation whose expression must be precisely balanced to maintain normal cellular growth while preventing serious pathologies that result from overexpression. Finally, trans-ribozymes maybe aptly suited to regulating expression within synthetic gene networks where transcript colocalization can be mediated via expression of the trans-ribozyme and target transcript from the same vector.

Introduction

Novel cellular behaviors can be generated by coupling engineered control systems to the cell's natural regulatory network [2]. However, the ability to run preprogrammed control algorithms depends on the facile construction of modular input / output interfaces between endogenous and synthetic systems. Currently, engineering gene circuits relies heavily on engineering protein-DNA interactions to control transcription. The rational design of protein-based control systems has been impeded by our limited understanding of how to engineer precise tertiary structures into proteins. Until it is possible to efficiently construct *de novo* connections between native and synthetic circuitry, the application of cellular programming strategies will be limited to a small set of input and output interfaces. As a result of these limitations, widespread application of exogenous control to cellular engineering applications remains unfeasible.

The modular and tunable nature of RNA makes it an ideal candidate to perform both sensing and actuation functions within an exogenous control system**.** Already researchers have demonstrated that RNA is capable of targeted gene knockdown through RNAi, trans-acting ribozymes, and antisense mechanisms [3-5]. However, the current technologies are limited in their ability to be integrated into robust, modular control systems. Off-target effects and poor efficiency hinder the effective application of antisense regulation of gene expression [5]. Hijacking the RNAi pathway has raised concerns that saturating the cellular machinery with heterologous substrates competes with the processing of endogenous RNAi substrates and can result in off-target or other undesired effects [6]. Additionally, RNAi is not universal throughout eukaryotic organisms such as *Saccharomyces cerevisiae*, a model organism for studying the

activation of various cancer pathways [7-9]. Ribozymes may offer an alternative gene regulatory mechanism upon which to build flexible, modular control systems.

Ribozymes are RNA molecules that catalyze a variety of chemical reactions such as self-cleavage or ligation [10]. The hammerhead ribozyme is comprised of three helical regions that converge on a highly conserved catalytic core of eleven nucleotides (nts) (Figure 4.1A) [11]. Cleavage is sequence-specific and targets a 5'-NUX-3' triplet, where N is any base, U is uracil, and X is any base except guanine. The optimal NUX for efficient and fast cleavage is GUC. Ribozyme cleavage is catalyzed when the 2' hydroxyl group from X directly 3' of the cleavage site is deprotonated. This nucleophile then attacks the scissile phosphate and, through a penta-coordinated trigonal bi-pyramidal transition state, produces a 5' and 3' product (Figure 4.1B) [12].

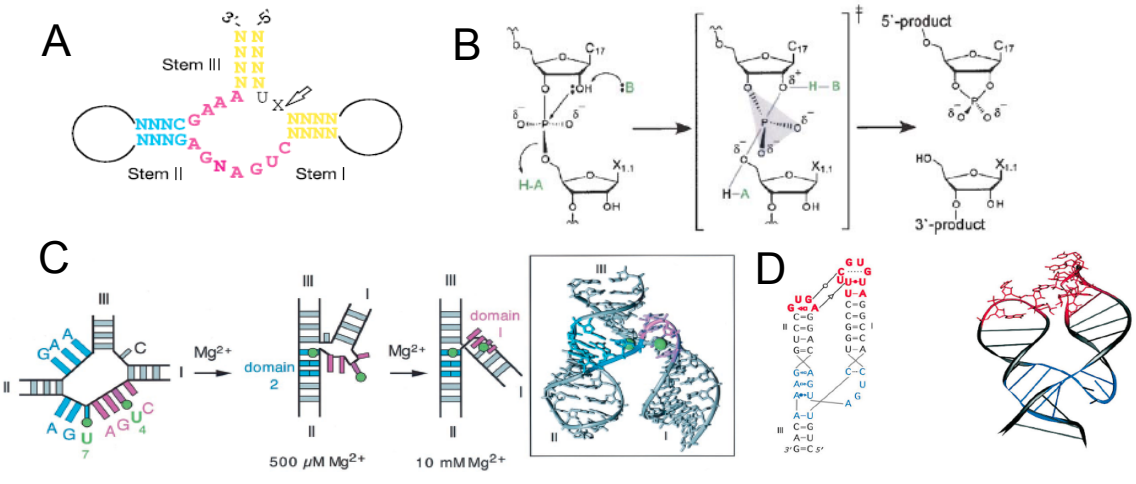


Figure 4.1. The hammerhead ribozyme. A) The catalytic core is shown in pink with the NUX sequence in black. The flanking helical regions are shown in yellow. Stem II is shown in blue. Black circles represent stem loops. Figure adapted from [13]. B) Ribozyme cleavage mechanism at the NUX triplet. Figure adapted from [12]. C) Folding of the ribozyme is thought to proceed through two magnesium binding events. Figure adapted from [14]. D) The "Y"-shaped ribozyme is thought to be stabilized by loop interactions between unpaired bases in the stem loops. Figure adapted from [11].

Folding of the ribozymes into an active conformation is postulated to proceed through dual divalent ion binding events (Figure 4.1C). A high affinity binding event occurs at 500 μM and orders the first set of tertiary interactions. The second low affinity addition of ion occurs at 10 mM and restructures the ribozyme stem orientations such that helix I folds away from helix III and interacts with helix II [14]. Ribozymes with a conserved catalytic core that do not maintain specific helical regions are called minimal ribozymes (mRzs). While at high divalent ion concentrations (10 mM) mRzs are active, at lower concentrations mRzs are effectively inert [11, 15]. Crystal structures of natural ribozymes depict a "Y"-shaped molecule that has two of the helical regions interacting as "kissing loops" (Figure 4.1D) [16]. Given the "kissing loop" interaction between the two helical regions, tertiary interactions between unpaired bases in the stem loops are proposed to stabilize the catalytically active conformation and obviate high divalent ion conditions. Researchers have demonstrated restored *in vitro* catalytic activity at biologically relevant divalent ion concentrations, between 100 and 300 μM, by reincorporating the loops into minimal ribozyme designs [11, 15, 17-20].

While the list of naturally occurring ribozymes is limited to cis-acting elements that perform intramolecular cleavage, synthetic trans-acting hammerhead ribozymes (thRzs) have been demonstrated that cleave intermolecular target sequences [4, 21, 22]. With their potential ability to target endogenously expressed transcripts, thRzs offer a significant advantage over their cis-acting counterparts, chRZs (Figure 4.2). Further, unlike cis regulatory elements, trans-acting elements like thRzs allow decoupling of ribozyme and target expression levels, offering the potential to tune knockdown via the independent modulation of thRz levels. While *in vitro* cleavage assays confirm that

trans-ribozymes cleave target transcripts specifically [23, 24], *in vivo* demonstrations of targeted gene regulation with trans-ribozymes are limited and restricted to higher eukaryotes [22, 25]. In addition, *in vivo* knockdown of gene expression remains hampered by low trans-ribozyme activity in the intracellular environment [20]. However, *in vivo* design rules have recently been specified and applied toward the construction of ribozymes that effectively regulate gene expression in cis through self-cleavage when placed in the 3' untranslated region (UTR) of target transcripts [26]. The application of the *in vivo* design rules that have been successfully implemented to convert cis-ribozymes into effective regulators of gene expression may be applied to trans-ribozymes to increase their efficacy as gene regulatory elements. Finally, while the simple knockdown of target genes is scientifically and clinically important, conditional regulation of gene products through integrated molecular control systems offers the potential to temporally and spatially tune therapies to restore proper function in diseased cells. Designing robust RNA-based control systems necessitates well-defined, modular actuators. By delineating the requirements for *in vitro* and *in vivo* activity we intend to examine the plasticity of the ribozyme domains requisite for targeting novel endogenous transcripts. Further, given the rules for constructing ligand-responsive chRz switches have recently been established [26], we evaluate the potential for constructing ligand-responsive RNA switches from the trans-acting ribozyme platform.

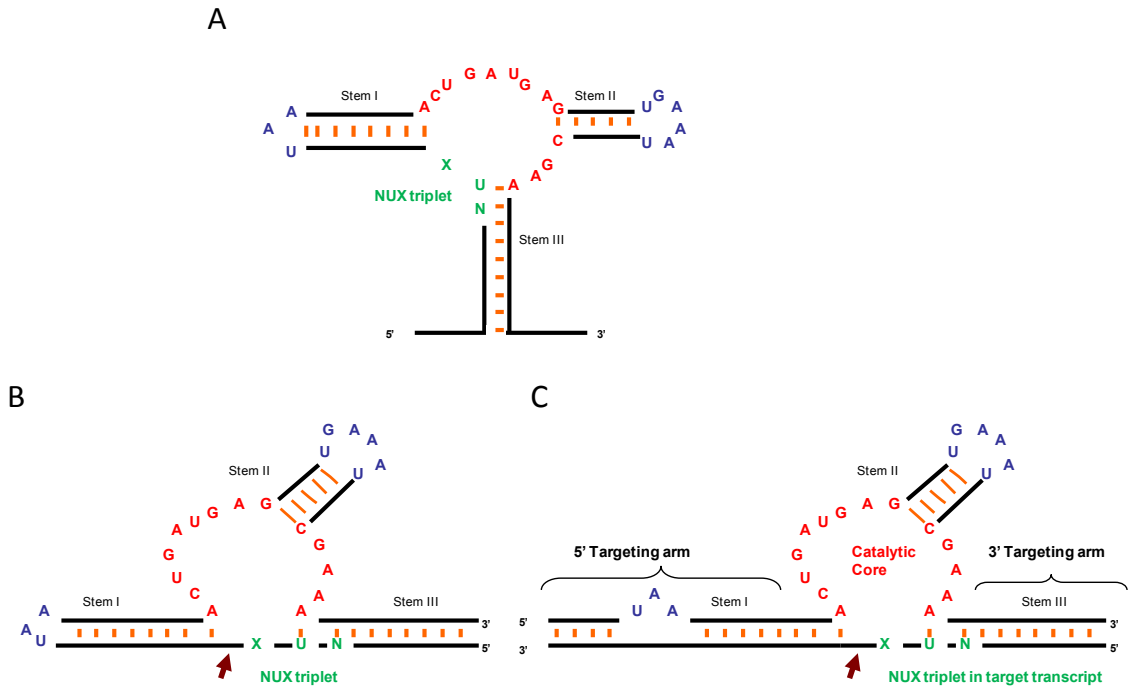


Figure 4.2. The anatomy of a cis-acting hammerhead ribozyme and trans-acting hammerhead ribozyme with target transcript. A. Canonical representation of cis-acting hammerhead ribozyme (chRz) with PLMVd loops incorporated into transcript via stem III. Stem I and stem II branch off catalytic core to give the eponymous hammerhead shape. **B.** A cis-acting hammerhead ribozyme with stem III rotated counter-clockwise 90^{o} with stem I fixed for simple comparison with trans-acting counterpart. **C.** Trans-acting hammerhead ribozyme (thRz) hybridized with target transcript. For all representations the catalytic core is shown in red and NUX triplet shown in green. Converting cis-acting ribozymes to trans-acting opens stem I from a closed loop. Preserving the loop (shown in blue) from the chRz require these nucleotides to be unpaired as a bulge. At low Mg^{+2} concentrations, the bulge in stem I and stem loop II (shown in blue), interact to stabilize the active conformation (Figure 1D). These interactions are found in natural chRzs and have been adapted for engineered thRzs.

In this work, we have taken the rules established for effective *in vivo* activity of chRzs and applied them to the development of thRzs. We demonstrate additional design constraints for efficient thRz cleavage in engineering the intramolecular binding event between the ribozyme and target transcript. In particular, we demonstrate that engineering the stem loops and the length of the targeting arms of the thRz is necessary to achieve efficient cleavage of the target transcript *in vitro*. Finally, we observe that thRz knockdown varies biphasically with target expression levels. Taken together these data

suggest that thRz knockdown is highly context-dependent and limited by system constraints that govern the intermolecular binding event.

Results

Initial thRz designs demonstrate *in vivo* limitations

The initial thRz designs were adapted from previous *in vitro* studies [20] to target a region in the transcript encoding yEGFP (yeast enhanced green fluorescent protein). We focused on developing thRzs that specifically target a fluorescent reporter protein (yEGFP) as an initial design goal as it allows us to quantify gene knockdown activity through fluorescence-based assays. In addition, yEGFP can be used to monitor various endogenous proteins via protein fusions, such that a trans-acting yEGFP ribozyme can be directly applied to regulate the expression of any coding sequence tagged with the target sequence from the yEGFP fluorescent protein. While the plasticity of the targeting arms is yet to be fully explored, we anticipate that the targeting arms will prove to be amenable to targeting a variety of sequences. Therefore, by modifying the targeting arms of the developed thRz construct we expect that we can adapt the system to target a wide array of cellular transcript.

The initial trans-ribozyme designs conformed to the following composition rules. First, the catalytic core was conserved to maintain activity (Figure 4.2C). Second, the target transcript contains a NUX triplet, where N is any base, U is uracil, and X is any base but guanine. Finally, for *in vivo* activity, the catalytically active "Y" shaped conformation must be formed. To achieve this requirement at physiological Mg^{2+} concentrations, the nucleotides in the bulges in stem I and stem II, are designed to

interact to stabilize the active conformation. These stem loop sequences are derived from those found in the peach latent mosaic viroid (PLMVd) ribozyme, a natural chRz [11]. Adopting these loops requires careful design considerations when selecting a target sequence. The bulge region must maintain unpaired nucleotides such that the targeted region in the target transcript does not have sequence complementarity to this region of the thRz. Additionally, canonical PLMVd ribozymes exhibit particular stem lengths between the stem I loop, the stem II loop, and the catalytic core, which were maintained in our designs. In stem I, 5 base pairs separate the bulge sequence from the catalytic core. In stem II, 4 base pairs separate the catalytic core and the stem loop sequence. Thus, canonical ribozymes will adopt the following nomenclature "X54" where X denotes the number of base pairs formed between the targeting arm sequence and the transcript 5' of the bulge (Figure 4.3).

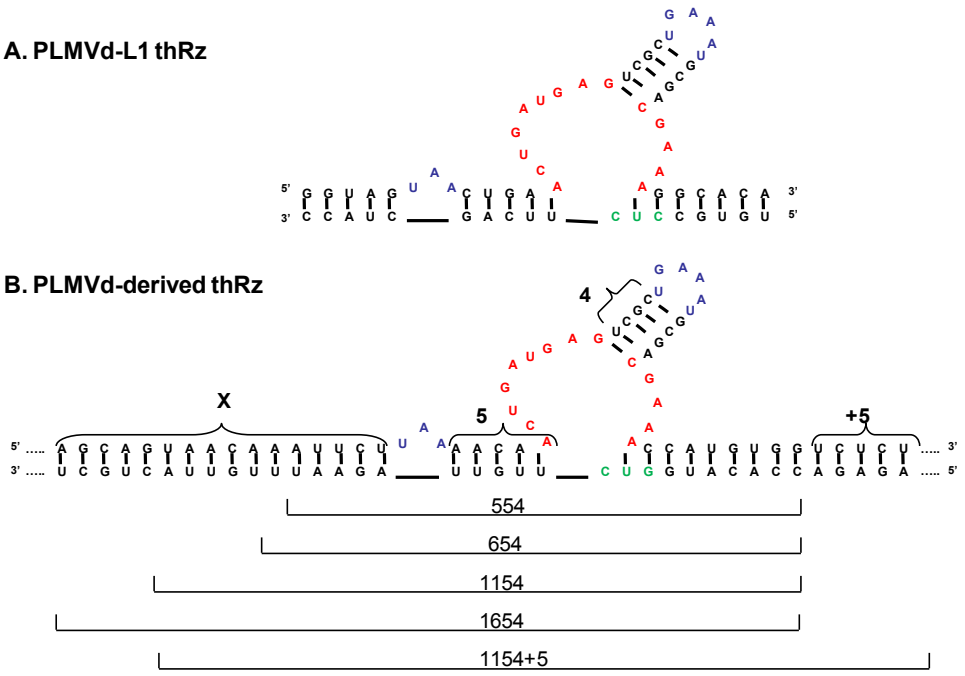


Figure 4.3. Various thRz designs. **A.** Previous *in vitro* studies focused on incorporating the PLMVd-derived stem loops into thRz targeted a segment of HIV1 mRNA, PLMVd-L1 thRz adapted from [20]. **B.** By redesigning the targeting arms, we created a series of PLMVd-derived thRzs targeting a sequence in yEGFP.

The initial thRz designs were characterized through *in vitro* cleavage assays in which separately transcribed ribozymes and radiolabeled target sequences were purified, combined, annealed, and incubated. Annealing was performed to allow transcripts to escape nonminimal free-energy structures potentially adopted during column purification. Endpoint cleavage efficiency was determined by comparing the amount of full-length to cleaved target sequence when visualized via PAGE gel and phosphorimaging analysis (see Methods). We modified our cleavage assay from previously reported methods to facilitate a more accurate correlation between *in vitro* results and *in vivo* activity. Specifically, we extended the length of the target transcript beyond the region that binds to the targeting arms. By including the peripheral transcript regions, we were attempting to more closely recapitulate the folding microenvironment of the full-length transcript for the region of interest. For example, the flanking sequences might interact with the target region to form secondary structures that might impede hybridization of the trans-ribozyme to the target region and thus cleavage. Accessibility of the target sequence is known to control the efficacy of thRz knockdown *in vivo* [27]. By targeting a RNA strand that more closely resembles *in vivo* transcripts, the *in vitro* cleavage assays are expected to reflect *in vivo* cleavage efficiencies more accurately. Following incubation, cleaved and noncleaved products were quantified through PAGE gel analysis. At 10 mM $MgCl_2$, the canonical trans-ribozymes, 554 and 654, cleave ~ 90% of the target (Figure 4.4A). The mRz with no loop structures cleaves ~ 50% of target. At Mg^{2+} concentrations comparable to physiological levels the activity of the mRz is completely abolished, whereas the canonical ribozymes maintain significant activity.

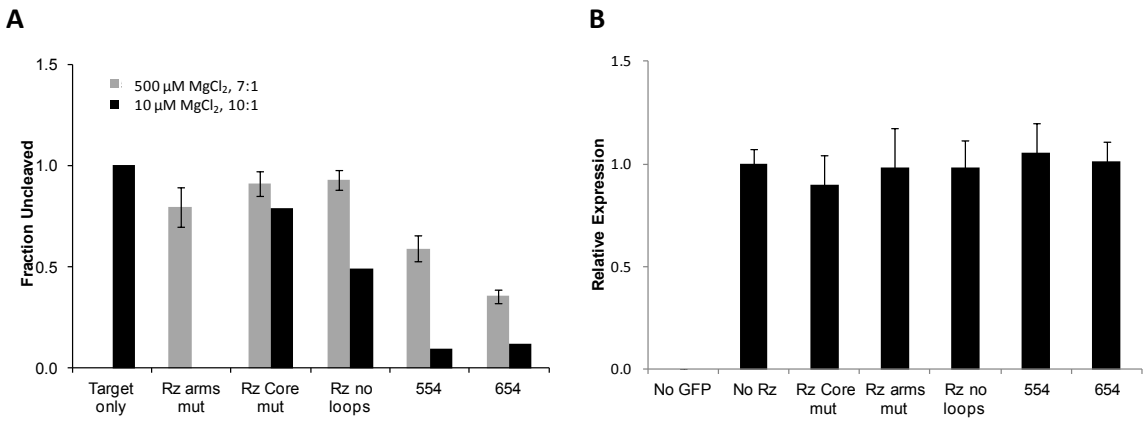


Figure 4.4. Initial ribozyme designs with pCS933. A. *In vitro* cleavage assays after 1 hr at various Mg^{+2} concentrations and ratios of ribozyme to target. **B.** *In vivo* fluorescence levels of yEGFP in various controls and constructs. Relative fluorescence levels were calculated by normalizing mean GFP levels of cells by the mean GFP level of the no ribozyme control. All fluorescence data are reported as the mean ± SD from at least three independent experiments.

We next investigated whether the cleavage activity observed *in vitro* translates to *in vivo* knockdown of target gene expression. In our system, thRzs were expressed from the strong constitutive TEF1 promoter on a high-copy plasmid, pCS933. The yEGFP target sequence fused to a PEST destabilization tag and placed under the control of the TEF1 promoter was integrated into the yeast chromosome. Cellular fluorescence of samples harboring trans-ribozyme expression plasmids and controls was analyzed by flow cytometry and the geometric means of gated cells were calculated (see Methods). Relative expression was calculated by subtracting the mean of the nonfluorescent control and normalizing by the no ribozyme control. The results indicate that despite exhibiting promising *in vitro* activities, these ribozymes did not yield observable knockdown of the target gene when expressed *in vivo* (Figure 4.4B). *In vitro* activity may not translate to *in vivo* knockdown for several reasons. First, our assays capture endpoint cleavage efficiency not kinetic rates of cleavage. Slow cleavage rates may prevent an observable change in gene expression. Alternatively, ribozyme colocalization and hybridization with

the target transcript *in vivo* may be limiting. Given that in the *in vitro* assay we anneal both strands together, the *in vitro* assay may not capture rate limitations with ribozyme-target hybridization and/or nonminimal global free energy states that inhibit the efficiency of hybridization.

Despite the limitations of the assay, the *in vitro* results from the canonical thRzs suggest that cleavage efficiency is limited by the ribozyme-target hybridization. At high Mg^{2+} concentrations the loop interactions should have little effect on the rate of cleavage since sufficient Mg^{2+} is present to stabilize the catalytically active conformation. Thus, observed differences between the mhRz's and the canonical thRzs' cleavage efficiencies at high Mg^{2+} may be a result of differing binding affinities. Increasing the length of the targeting arm is expected to increase the affinity of the ribozyme for its target, facilitating ribozyme-target hybridization. The targeting arms of the mRz are 4 nts and 5 nts shorter than 554 and 654, respectively. At 10 mM, this difference in targeting arm length yields a 40% difference in cleavage efficiency between the mRz and the canonical ribozymes (Figure 4.4A). At 500 μM, the difference in the length of the targeting arms is magnified. Ribozyme 654, which only differs from 554 by a single nucleotide in the targeting arms, is ~ 25% more efficient. These results led us to speculate that increasing the length of the targeting arms may significantly increase the *in vivo* gene-regulatory efficiency of a trans-ribozyme element. In addition to increasing ribozyme affinity for target, increasing the targeting arm length should favor proper ribozyme binding to the target.

Optimization of thRz targeting arms leads to fast, efficient cleavage at physiological $MgCl_2$ concentrations

To test whether increasing the targeting arms improves ribozyme cleavage efficiency, we constructed three new thRz designs that incorporated longer targeting arms (Figure 4.3B). One design (1154) increased the length of the 5' arm by 5 nts. The second (1154+5) and third (1654) designs built on the first by adding 5 nts to the 3' and 5 nts to the 5' end, respectively (Figure 4.3B). When the resulting ribozyme sequences were folded in RNAstructure 4.3 with the 137 nt target sequence, the proper ribozyme-binding-to-target structure dominated the energy landscape as the minimum free energy (MFE) structure and was the only structure found in the 20 lowest free energy structures (Supplementary Figures 4.1-3). When tested in the *in vitro* ribozyme cleavage end point assays the canonical trans-ribozymes with increased targeting arm length, resulted in greater than 90% cleavage of the target transcript at physiological Mg^{2+} concentrations (Figure 4.5A). Furthermore, the thRzs with increased targeting arm length (1154, 1154+5, 1654) exhibit greater cleavage efficiencies at low Mg^{2+} concentrations than the previous canonical thRz design (554, 654) at higher Mg^{2+} concentrations. These results support that the limiting step in the cleavage reaction is the intramolecular binding event between the thRz and target sequences. Formation of the catalytically active conformation appears to occur readily once the target is found as indicated by the difference between 654 and 1154. However, increased targeting arm lengths may also contribute to increased cleavage efficiency by increasing the stability of the active tertiary conformation as well as the formation of the correct secondary structure. Additionally, the *in vitro* cleavage assays demonstrate that lengthening the targeting arms facilitates

very rapid cleavage of target transcript (Figure 4.5A). Within 10 minutes, all of the redesigned thRzs reach 97% cleavage efficiency or higher, while even after an hour thRzs with shorter arms showed 90% or less cleavage efficiency. For *in vivo* application it is important that the ribozymes act on biologically relevant time scales (~several minutes). While previous *in vitro* studies have demonstrated fast and efficient thRz cleavage at higher Mg^{2+} concentrations (1 mM), these results suggest that full-kinetic evaluation will prove these thRzs to be highly efficient at 500 μM (on the order of minutes).

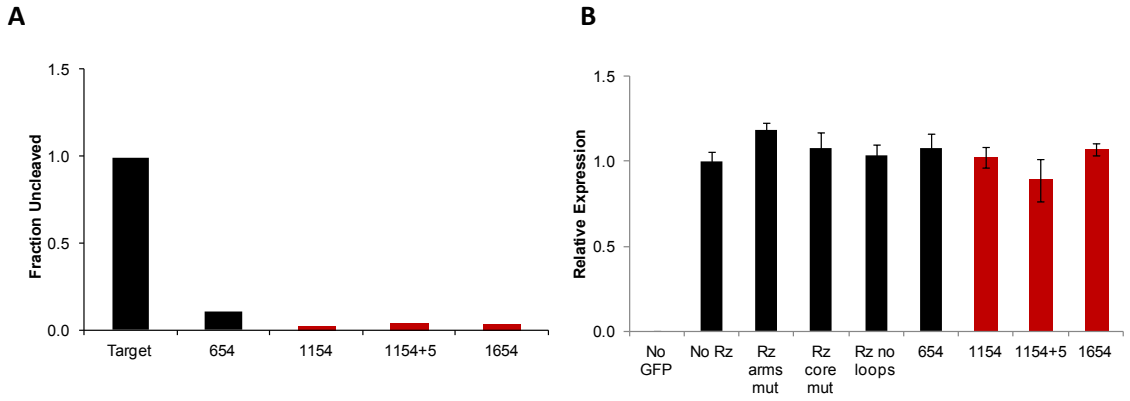


Figure 4.5. Initial ribozyme designs with extended targeting arms show improved *in vitro* efficiency but fail to knockdown expression *in vivo*. A. *In vitro* cleavage assay after 10 minutes at 500 uM Mg^{+2}. **B.** *In vivo* fluorescence levels of yEGFP in various controls and constructs. Relative fluorescence levels were calculated by normalizing mean GFP levels of cells by the mean GFP level of the no ribozyme control. All fluorescence data are reported as the mean ± SD from at least three independent experiments.

The gene-regulatory activities of the thRzs with lengthened targeting arms were assayed *in vivo* as described previously from the pCS933 expression system. However, despite very promising *in vitro* results, these optimized trans-ribozymes did not yield observable knockdown of the target gene when assayed in the *in vivo* expression system (Figure 4.5B). We suspect that topological constraints on the ribozyme transcript may hinder ribozyme-target binding, potentially explaining the lack of observable cleavage activity of these ribozymes *in vivo.* The ribozymes are expressed from a Pol II promoter,

such that during transcription, a modified nucleotide called the 5' cap is added to the 5' end of the ribozyme-encoding mRNAs. As protein factors accumulate on the transcript, the cap binds to the poly-A tail and circularizes the transcript. We postulated that in the ribonucleoprotein particle (RNP) ribozymes may be inhibited from binding to the target transcript by the topological constraints imposed by circularization of the transcript and binding of proteins to the transcript. For example, ribosome loading onto the transcript through the 5' cap structure may occlude the targeting arms from being free to interact with the target transcript.

Implementation of a novel expression system improves the knockdown efficiency of the trans-ribozyme elements

To test whether knockdown could be improved by preventing topological constraints imposed during transcript circularization, we designed and constructed an expression cassette that would allow the thRz to be separated from the RNP. Specifically, two cis-acting hammerhead ribozymes (chRzs) were placed on either side of the thRz element (into pCS933 at AvrII and SacII restriction sites) (Figure 4.5). This new thRz expression cassette contains two chRz, previously shown to be highly efficient at self-cleaving *in vivo* [11, 15], separated by two unique restriction sites, SacI and SphI, for the rapid cloning of the thRz between these two elements (pCS975). The chRzs are expected to excise the thRz from the RNP through self-cleavage, resulting in removal of the transcript tails containing the 5' cap and the poly-A tail. Removal of the 5' cap and poly-A tail will also prevent ribosomes from loading onto the transcript. Additionally, by isolating the liberated transcript from promoter- and terminator-specific transcript tails,

the chRz processing cassette is expected to facilitate thRz portability across a range of expression systems.

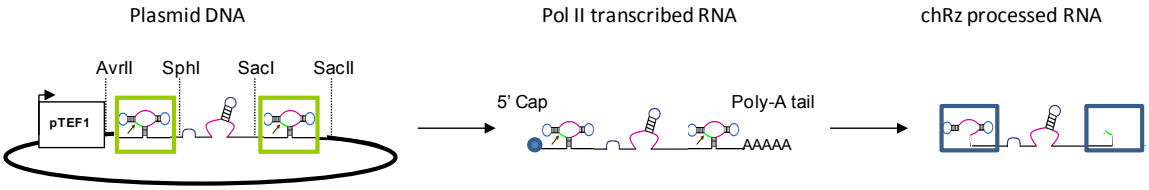


Figure 4.6. Redesigned expression system. The chRz processing cassette was cloned into pCS933 at AvrII and SacII sites to create pCS975. thRz were subsequently cloned between SphI and SacI. pCS975 expresses the trans-ribozyme with flanking cis-hRz (boxes in green) on either side of the thRz transcript designed. Following transcription the chRZs trim the transcript tails, removing the 5' cap and poly-A tail to yield the chRz-processed RNA. Remaining chRz tails shown in blue boxes.

To test *in vitro* activity of the ribozymes in the context of the new expression system, the ribozymes were transcribed from pCS975 containing the chRz cassette along with promoter- and terminator-tails as described previously. *In vitro* cleavage assays confirmed that the chRzs cleave with near-perfect efficiency during transcription as expected given the high Mg^{2+} concentration of the T7 transcription buffer (data not shown). The resulting thRzs with chRz tails were incubated with target at various Mg^{2+} concentrations. At physiological Mg^{2+} concentrations, the thRzs in the context of the chRz-processing cassette show significant sensitivity to targeting arm length. Activity is completely abolished for the thRz with shorter targeting arms (554, 654), but restored for the thRzs with longer targeting arms (Supplementary Figure 4.4A and Figure 4.7A). Overall, the thRzs flanked by the chRzs demonstrated lower cleavage activity compared to the original expression system while maintaining similar trends in cleavage efficiencies. These data suggest that the chRz tails that remain on the trans-ribozyme transcript following cis-cleavage may interfere with thRz activity especially when the targeting arms are shorter. By investigating the predicted structures of ribozymes folded

with target transcripts in RNAstructure 4.3, we determined that the chRzs' tails are predicted to interfere with thRz binding to the target by preferentially promoting formation of Watson-Crick bonds between the target sequence and the catalytic core (Supplementary Figures 4.5–4.7). The resulting competition between the core region and targeting arms for binding to the target sequence may result in the observed decrease in the activity of the trans-ribozyme elements. This effect may be magnified at low Mg^{2+} concentrations, where proper folding of the ribozyme is not aided by increased levels of this ion.

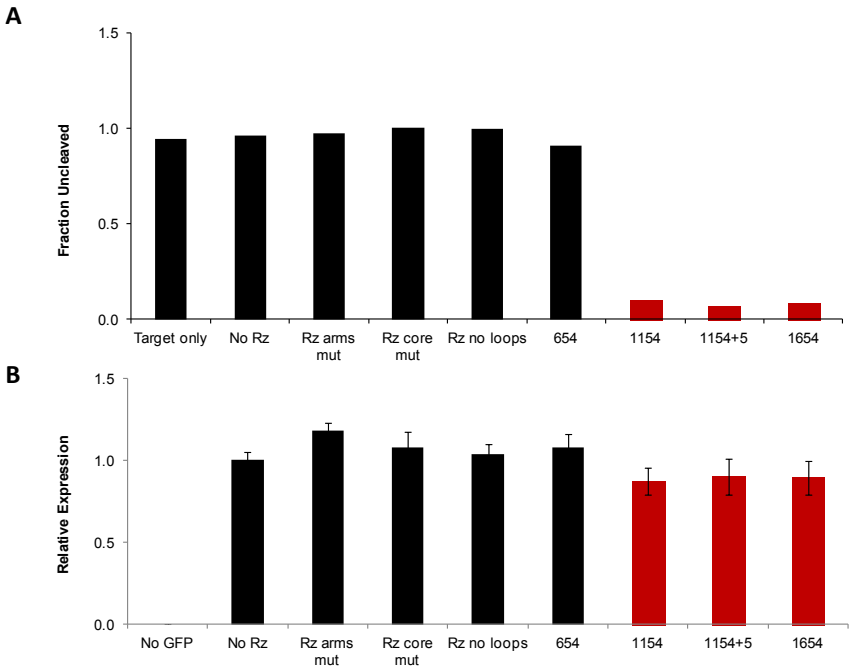


Figure 4.7. Improved ribozymes with chRz-processing expression cassette show only modest *in vivo* knockdown despite significant improvement *in vitro*. A. *In vitro* cleavage assays after 1 hr at 500 uM Mg^{+2}. **B.** *In vivo* fluorescence levels of yEGFP in various controls and constructs. Relative fluorescence levels were calculated by normalizing mean GFP levels of cells by the mean GFP level of the no ribozyme control. All fluorescence data are reported as the mean ± SD from at least three independent experiments.

In vivo assays of ribozyme gene-regulatory activity were performed as described previously to determine if the chRz-processing cassette improved ribozyme-mediated knockdown. *In vivo* expression of the ribozymes from the redesigned plasmid system

demonstrated minimal knockdown of gene expression (~ 10%) (Figure 4.7B). Given that we had established efficient catalytic activity of our designs at low Mg^{2+} concentrations (through the *in vitro* cleavage assays) and modified the design of the expression construct to eliminate potential interference from the RNP and ribosome loading via the chRz processing cassette, we suspected that the *in vivo* gene-regulatory activity of the trans-ribozyme system may be limited by efficient localization and thus hybridization of the trans-ribozyme element to the target transcript in the dense cellular milieu. Additionally, it is possible that by excising the thRzs from the RNP stability of the ribozyme transcript was compromised. Destabilized transcripts would be expected to have short residence times, potentially shorter than the time necessary to locate and hybridize to a target transcript.

Varying target expression levels indicates a biphasic relationship between target concentration and *in vivo* knockdown activity of the trans-ribozyme elements

To determine whether *in vivo* knockdown could be improved by modulating the expression of the target, we modified the *in vivo* assay system such that the level of target transcript could be tuned relative to the level of ribozyme being expressed. It is expected that increasing the level of target transcript will increase the probability of trans-ribozyme binding to the target, thereby potentially increasing ribozyme-mediated knockdown within a particular range of target transcript. Above this range, as ribozymes become saturated by target, knockdown would be expected to taper off and ultimately become unobservable as the rate of binding and cleaving the target transcript is dwarfed by the rate of target expression. Due to these competing effects of target transcript levels, there

may be an optimal level of target expression at which to observe knockdown. Given that we anticipated that ribozyme saturation could be an issue, the original and improved thRz expression systems were constructed on a high-copy plasmid with a strong TEF1 promoter to maximize thRz expression. To further increase the ratio of ribozyme to target we modified the target expression system to allow titration of the target transcript over a lower range of expression levels. An expression construct encoding the yEGFP gene regulated from a galactose-inducible promoter, pGAL1, was integrated into the yeast chromosome to make strain CSY341. Integrating the yEGFP cassette at the GAL2 locus, converted the all-or-none response of the galactose-inducible promoter to a linear, homogenous response to galactose as was previously shown [28]. With this modified expression system the target transcript levels increase as a function of increasing galactose concentrations. The previous target expression strain CSY132 has yEGFP integrated with pTEF1 and a PEST destabilization tag. The TEF1 promoter is more than 10 times stronger than the pGAL1 promoter induced with 2% galactose for 6 hours (Supplementary Figure 4.8).

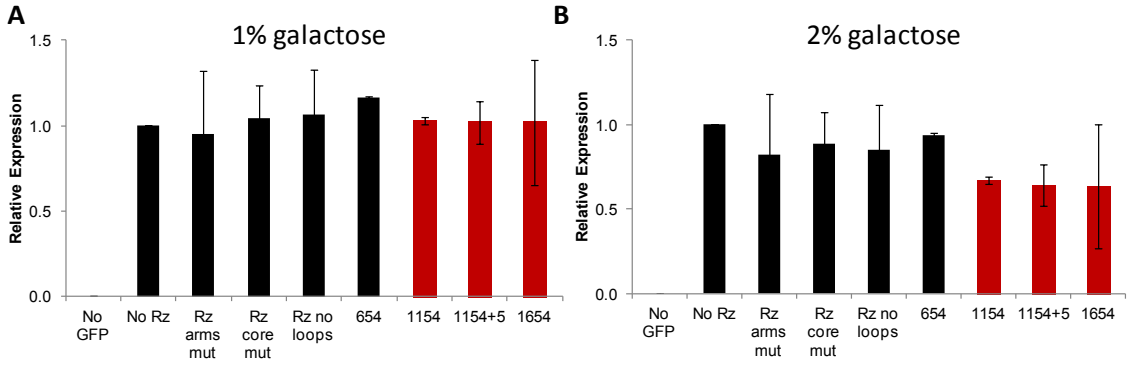


Figure 4.8. Controlling target expression via the galactose-inducible promoter demonstrates that knockdown increases at higher expression levels of target transcript. A. *In vivo* fluorescence levels of yEGFP in various controls and constructs for thRz in chRz-processing expression cassette (pCS975) at 1% galactose and **B**. 2% galactose. Target transcript levels increase with increasing galactose levels. Relative fluorescence levels were calculated by normalizing mean GFP levels of cells by the mean GFP level of the

no ribozyme control. All fluorescence data are reported as the mean ± SD from at least three independent experiments.

We tested the thRz with and without the chRz processing cassette expression system, pCS933 and pCS975, respectively, for *in vivo* knockdown of different levels of target expression. Cells were grown overnight in noninducing, nonrepressing conditions and induced at 1% or 2% galactose for 6 hours before assaying yEGFP via flow cytometry (Materials and Methods). Relative expression was calculated by normalizing to the mean fluorescence of the no ribozyme control at the corresponding galactose concentration. Raw fluorescence means increase on average 3-fold from 1% to 2% galactose. At 1% galactose, ribozyme-mediated knockdown activity is not observed for the thRzs expressed with or without the chRz processing cassette (Figure 4.8A and Supplementary Figure 4.9A, respectively). However, both trans-ribozyme expression systems exhibit improved knockdown ranging from 16%–35% when target gene expression is increased by adding 2% galactose to the cells (Figure 4.8B and Supplementary Figure 4.9B). This data supports the hypothesis that mediating the ribozyme-target binding event is the limiting step *in vivo.*

In evaluating the efficacy of the thRzs across a range of target expression levels, the results indicate that increasing expression levels by ~ 3-fold (from 1% to 2% galactose) increases knockdown from 0% to 35%. However, in the pTEF1-yEGFP-PEST target expression system where expression levels are estimated to be an order of magnitude higher than in the galactose-titratable CYS341 system at 2% galactose, very minimal knockdown is observed (Figure 4.7B). Taken together, these data suggest that ribozyme-mediated knockdown may exhibit a biphasic response to target transcript levels.

Discussion

We have shown that thRzs can be engineered for improved catalytic activity *in vitro* that translates to *in vivo* knockdown dependent on target transcript expression. Through the rational design of the targeting arms aided by computational models of RNA secondary structure, *in vitro* thRz activity at physiologically relevant Mg^{2+} concentrations was significantly increased. Additionally, by modifying our *in vitro* cleavage assay such that the target sequence more accurately reflects the sequence and associated structure targeted *in vivo*, we developed a more instructive *in vitro* cleavage assay for evaluating the potential of thRz designs to efficiently cleave target transcripts at physiologically relevant Mg^{2+} concentrations. Results from the improved *in vitro* assays indicate that increasing targeting arm length improves thRz cleavage efficiency. Further, at these lower Mg^{2+} concentrations, our *in vitro* results support that the designed thRzs efficiently cleave the target on biologically relevant timescales. We observed that implementation of a chRz processing cassette that releases the thRz from the RNP facilitates modest knockdown of the target *in vivo*, an improvement over our initial expression system. Finally, varying target expression levels suggests that ribozyme-mediated knockdown *in vivo* is highly sensitive to target expression levels and responds biphasically to increasing target expression.

To capture binding constraints imposed by the structures formed by the sequences flanking the target sequence on the target transcript, we modified our *in* vitro assay to analyze cleavage of a longer target sequence. Previous *in vitro* analysis of trans-ribozyme efficiency evaluated ribozymes targeting transcripts that were 18 nts in length, such that

the targeting arms hybridized to all but one nt of the sequence [20]. In our modified assay we transcribed 137 nts of the target sequence with 86 and 51 nts on either side of the cleavage site. We anticipated that a target of this length would capture local interactions of the target sequence with the flanking sequence, thus providing us with an assay that captures potential limitations to ribozyme hybridization and cleavage that may exist in the *in vivo* conditions. While previous assays utilizing short target sequences showed that canonical ribozymes with short targeting arms rapidly achieve high cleavage efficiencies (> 90% in 5 minutes) at low Mg^{2+} concentrations [20], our results with the longer target transcript demonstrate that increased targeting arm length was required for similarly efficient cleavage. Further, only thRzs with elongated arms achieved observable knockdown *in vivo*, suggesting that assaying thRz with longer target transcripts improves the translation of cleavage efficiency observed *in vitro* to knockdown observed *in vivo*

Despite improving the ability of the *in vitro* assay to capture the impact of longer target transcripts on cleavage efficiency, the assay could be further improved to mimic the additional constraints imposed by the *in vivo* environment on thRz binding and cleavage. Assays were performed at 50 mM Tris-Cl pH 7.0, 100 mM NaCl at different $MgCl_2$ concentrations as described in previous work [20]. This reaction buffer is unlikely to model the crowded cellular environment. Supplementing the buffer with PEG and/or BSA could potential provide a closer model of the dense cellular milieu in which ribozymes and target transcripts must hybridize. Additionally, since we performed our assays by annealing the ribozyme and target sequences together following purification, the assay could be improved by separately annealing the sequences. *In vivo* ribozymes and target transcripts are transcribed and folded independently, which is likely to result in

important intramolecular structures that can influence ribozyme-target binding. To more accurately represent potential limitations to hybridization of the thRz and target transcript that arise in the intracellular folding environment, individual annealing steps to remove entrapped, nonminimal energy structures induced during the purification step can be performed prior to combining the transcripts or potentially dispensed with all together. While it is desirable to cotranscribe ribozymes with the target, the composition of T7 transcription buffer precludes assessing cleavage efficiency at low Mg^{2+} concentrations, an important requirement for accurately modeling the effect of the intracellular environment on cleavage efficiency.

We demonstrated that the chRz-processing expression system modestly improves knockdown over the original system at high target expression levels. However, no significant difference is observed at low target expression between the two ribozyme expression systems examined in this study. Because the chRz-processing cassette removes the 5' cap and poly-A tail that acts to stabilize transcripts, this expression system may reduce the half-life of expressed thRzs *in vivo.* One explanation for the observed behavior is that at high target expression levels, removing the topological constraints imposed by the RNP may improve ribozyme hybridization rates such that a reduced half-life is less consequential. However, at lower target expression levels, where the average time to bind to a target sequence is expected to increase, thRz half-life may more significantly impact knockdown efficiency and erase efficiencies gained from removing topological constraints associated with the RNP. The expression system might be further improved by the addition of strong hairpin elements at the 5' and 3' ends of the thRz. Such hairpins can increase the residence time of transcripts by protecting against

exonuclease degradation [29, 30]. In addition, spacer regions can be designed to separate the hairpins from the thRz, which may allow for greater insulation of the thRz targeting arms from any bound nucleases or other proteins. Expression systems that can prevent proteins from occluding the thRz targeting arms and increase thRz half-life may result in improved knockdown efficiency.

Varying the levels of the target indicated that ribozyme-mediated knockdown exhibits a biphasic dependence on target expression levels. Specifically, at very low expression levels, no knockdown is observed. Increased expression (3-fold) above this low level results in an increase in the observed knockdown, suggesting that *in vivo* the intramolecular binding event may be limiting knockdown at low target concentrations. However, increasing expression by an order of magnitude reduces observable knockdown potentially due to the target transcript saturating the available ribozyme pool. Probing the range and threshold of target expression for which knockdown is observed could provide a better understanding of the optimal range of target expression levels for thRzs. Delineating this response curve to target expression levels would be useful in identifying potential systems in which trans-acting ribozymes may be effectively applied as regulators of gene expression. The range and threshold of target expression permitting thRz-mediated knockdown are likely to be sequence-specific, thus it would be important to identify a high impact set of targets for ribozyme-mediated knockdown. We did not further probe the threshold and range in our system due to a few system limitations. One drawback of the galactose-titratable system is the inherent noisiness of the galactose promoter. For some samples such as pCS975 1654 (Figure 4.8B), the replicate variability did not allow for significant differentiation from the controls given the error associated

with the assay and system. A more comprehensive and precise study of the relationship between knockdown and target expression levels could be performed by assaying thRz gene-regulatory activities in strains in which the target is expressed from constitutive promoters of varying strength and integrated into the chromosome.

While we have addressed several potential limitations to ribozyme-mediated knockdown, there remain several additional avenues that should be explored to further improve knockdown efficiency. Previous *in vivo* work with thRzs has shown success in targeting the HIV-1 transcript, which is localized to the nucleolus [4]. Researchers used an expression system that localizes thRzs to the nucleolus in order to achieve substantial knockdown of HIV-1 transcripts. Thus, developing a mechanism to localize ribozymes to the target transcript may broadly improve knockdown across a range of target transcript levels. While for many endogenous genetic targets it is not possible to colocalize sequences to cellular substructures, this strategy may be effective for particular systems in which transcripts localize to specific parts of the cell. Additionally, the localization constraints for transcripts may position trans-ribozymes as effective regulators within the context of synthetic gene networks where transcript colocalization can be mediated via expression of the trans-ribozyme and target transcript from the same vector. One final consideration in developing ribozymes as effective regulators of gene expression is the accessibility of the target sequence. Several studies have examined the accessibility of sequences within transcripts for binding to nucleic acids and developed algorithms that predict sequences that exhibit enhanced binding accessibility [31-33]. The utilization of such algorithms may improve the design of trans-acting ribozymes by identifying accessible target sequences.

While the development of a trans-ribozyme platform represents an opportunity to regulate the expression of target transcripts without the need to modify existing, endogenous cellular components, such as binding sites, promoters, repressors, and other cis-acting regulatory elements, or saturating the natural regulatory RNAi pathway, the limited activity observed from these systems *in vivo* constrain the potential applications of this system. Based on these studies, systems with moderately high levels of target transcript that require only modest levels of knockdown are potential candidates for thRz regulation. For example, diseases such as cancer and Alzheimer's are characterized by aberrantly high mRNA levels of particular genes [7, 34, 35]. In Her2 positive cancer cells, overexpression of the Her2 receptor leads to increased cellular growth via upregulation of the signaling pathway [35]. High Her2 expression negatively correlates with survival rates in endometrial cancer [34]. Monoclonal antibodies that target the Her2 receptor reduce tumor proliferation and increase survival rates presumably by binding to Her2 and blocking signaling. Further, siRNA knockdown of Her2 sensitizes cells to monoclonal antibody therapy [34]. However, the systemic administration of these antibodies has resulted in lethal side-effects such as cardiac dysfunction [36]. Broad distribution of these antibodies particularly to cardiac tissue is suspected to reduce signaling critical to preventing cardiac myopathy [35]. To improve therapies for Her2 positive cancers, a thRz targeting the overexpressed Her2 transcripts could provide a therapy targeted to diseased cells. At high Her2 levels, a modest yet potentially significant knockdown of the target transcript could reduce aberrant signaling and additionally sensitize cells to monoclonal therapy. Further, such thRz therapies could

prevent lethal side effects by remaining inert in the case of lower expression levels characteristic of healthy cells.

There remain several challenges in developing such a therapy. First, the expression levels of target transcripts in healthy and diseased cells would need to be determined and evaluated. To reduce expression in diseased cells, expression must be above the requisite threshold for trans-ribozyme-mediated knockdown, but not so high as to saturate the ribozymes present in the cell. Additionally, to optimally differentiate healthy and diseased cells and limit side-effects, expression levels for each cell type should fall below and above the minimum threshold, respectively. Second, utilizing the principles elucidated above, various thRzs would need to be designed to the target transcripts. Utilizing previously mentioned algorithms may aid in the identification of accessible target sites within the transcript. Thirdly, the threshold for knockdown would need to be evaluated to determine if significant knockdown of the targets can be achieved in the diseased cells and if there are any potential effects on healthy cells. It would be useful to determine if the threshold of target transcript knockdown can be tuned via the design of the targeting arm lengths, such that transcripts across a broad range of concentrations can be targeted, increasing the range of potential systems to which these thRzs could be applied. Additionally, differences in metabolism, particularly those affecting transcript degradation and translation rates, may alter the *in vivo* efficiency of thRzs between organisms. This consideration may explain why previous examples of thRzs *in vivo* knockdown have been mostly limited to higher eukaryotes [22, 25]. Finally, ligand-mediated control of ribozyme activity would be a useful tool in directing thRzs as a therapeutic agent. However, the limited efficiency observed in our *in vivo* test

system implies that the current designs would not be amenable to switch construction. While the rules for construction of ligand-responsive switches from chRzs are expected to be readily transferable to thRzs, constructing switches requires appending sequences to the thRz stem II loops. In chRzs, appending sequences to the stem II loops has been shown to reduce cleavage efficiency compared to the nonswitch control [26]. For the successful construction of ligand-responsive thRz switches, ribozymes are anticipated to require higher *in vivo* activity than those developed in this study to target yEGFP.

While significant progress has been made in understanding the catalytic behavior of trans-acting ribozymes *in vitro* in this work and others [17, 18, 20], this progress has not yet been translated beyond modest *in vivo* gene regulation. Given the success of cis-acting ribozymes *in vivo,* the key limitation to efficacy of trans-acting ribozymes appears to be facilitating the intermolecular hybridization interaction between the trans-ribozyme and target mRNA sequences. While this limits the potential application space of trans-acting ribozymes, it may position these devices as uniquely targeted regulators of gene expression for systems requiring modest knockdown of targets that are overexpressed and near quiescence at normal low levels. Additionally, trans-ribozymes may be useful control tools within synthetic circuits where there exists greater design flexibility for mediating transcript colocalization. Future work on these devices should focus on identifying potential systems for application and designing thRz to function in these systems as biochemical controllers of gene expression.

Materials and methods

Plasmid and ribozyme construction

Standard molecular biology cloning techniques were used to construct all plasmids [37]. DNA synthesis was performed by Integrated DNA Technologies (Coralville, IA). All enzymes, including restriction enzymes and ligases, were obtained through New England Biolabs (Ipswich, MA). Ligation products were electroporated with a GenePulser XCell (Bio-Rad, Hercules, CA) into an *E. coli* DH10B strain (Invitrogen, Carlsbad, CA), where cells harboring cloned plasmids were maintained in Luria-Bertani media containing 50 mg/ml ampicillin (EMD Chemicals). All cloned constructs were sequence verified by Laragen Inc (Santa Monica, CA).

The thRz expression constructs, pCS933 and pCS975, were constructed from pCS346 (Supplementary Figure 4.10). pCS346 bears two sets of TEF1 promoters and CYC1 terminators with intervening unique restriction sites. Briefly, pCS933 was constructed by cloning a version of RFP (tdimer2), which served as a transformation control signal, into the SalI and NotI restriction sites downstream of the second TEF1 promoter in pCS346. Site-directed PCR mutagenesis was performed to remove the two SacII sites from tdimer2, preserving SacII as a unique restriction site using the QuikChange II Site-directed mutagenesis kit (Agilent, Santa Clara, CA) according to manufacturer instructions. The pCS933 engineered ribozyme constructs were generated by cloning the appropriate thRz constructs into the unique restriction sites, AvrII and SacII, downstream of the first TEF1 promoter (see Supplementary Table 4.1 and Supplementary Table 4.2). The thRz with flanking chRzs plasmid, pCS975, was constructed by cloning a cassette containing two chRz with two intervening unique

restriction sites, SphI and SacI, into AvrII and SacII using the primers Cis-Cis.sTRSV.FWD and Cis-Cis.sTRSV.REV. The pCS975 engineered ribozyme constructs were generated by cloning the appropriate thRz constructs into the SphI and SacI sites. All oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, IA). Cloned plasmids were transformed via electroporation into an electrocompetent *Escherichia coli* strain, DH10B (Invitrogen, Carlsbad, CA), and all cloned ribozyme constructs were confirmed by sequencing (Laragen, Los Angeles, CA).

Confirmed plasmid constructs were transformed into CSY132, a *S. cerevisiae* W303 strain harboring a chromosomally integrated pTEF1-yEGFP-PEST target construct (*MATα his3-11,15 trp1-1 leu2-3 ura3-1 ade2-1, pTEF1-yEGFP-PEST*) or CSY341, a *S. cerevisiae* W303 strain harboring a chromosomally integrated pGAL1-yEGFP target construct (*MATα his3-11,15 trp1-1 leu2-3 ura3-1 ade2-1,* Δgal2, *pGAL1-yEGFP*) using a standard lithium acetate procedure (Supplementary Table 4.3) [38]. To construct CSY341, the pGAL1-yEGFP cassette was PCR amplified from pCS1340 (Supplementary Figure 4.11) using primers GAL2ko.fwd and GAL2ko.rev to integrate the cassette at the GAL2 locus (Supplementary Table 4.4). Yeast strain CSY3 was transformed as described previously with 12 μg of gel purified PCR product and plated on G418 plates to build yeast strain CSY341.

In vitro endpoint cleavage assays

All ribozymes were PCR amplified from their plasmids beginning at 23 nts upstream from the TEF1 promoter using primers T7Rbz 5'-23.FWD and T7Rbz 3'.REV (Supplementary Table 4.5). A 137 nt region of the target yEGFP sequence was amplified

by PCR using Target.yEGFP.FWD and Target.yEGFP.REV, where the forward primer in each of these amplification reactions harbors the T7 polymerase sequence at its 5' end. PCR products were transcribed using an Ampliscribe T7 kit (Epicentre Technologies, Madison, WI) according to manufacturer's instructions. Transcription reactions were DNaseI treated for 15 min. at 37^{o}C and purified into reaction buffer (50 mM Tris-Cl pH 7.0, 100 mM NaCl, and the specified $MgCl_2$ concentration) through a NucAway column (Ambion, Foster City, CA) following the manufacturer's instructions. The target transcript was radiolabeled by modifying the reaction solution to include 2 μl of [α-32P]-GTP and 1 μl GTP. Cleavage reactions were performed in a 20 μl total volume at 50 mM Tris-Cl pH 7.0, 100 mM NaCl, and the specified $MgCl_2$ concentration. The ratio of ribozyme to target was 10:1 and 7:1 was specified by varying the volume of purified ribozyme and target RNA added to reactions. The ribozyme and target sequence were annealed in reaction buffer by incubating at 95°C for 5 min and cooling at room temperature for 15 minutes. Reactions were run at 37°C for either 10 or 60 minutes and quenched with the addition of RNA loading buffer II (Ambion). Prior to loading on the gel, reaction products were heated at 65°C for 5 minutes and chilled at 4°C for 5 minutes. The reaction products were separated on a 6% denaturing PAGE gel, dried, and visualized on a FX phosphorimager (Bio-Rad, Hercules, CA). The RNA decade ladder (Ambion Foster City, CA) was used as a size marker. The background-subtracted intensity of full-length target (137 nt) and cleaved products (~50 and 85 nts) were quantified for each sample using the Quantity One software package (Bio-Rad). The fraction of RNA uncleaved was determined by dividing the intensity of full-length target by total intensity.

In vivo ribozyme characterization assays

S. cerevisiae cells (W303) harboring the appropriate plasmids were grown in synthetic complete medium supplemented with an appropriate dropout solution and sugar [2% (wt/vol) dextrose] overnight at 30°C. For galactose titration, 2% raffinose and 1% sucrose were substituted for dextrose. Overnight cultures were back-diluted into fresh medium to an optical density at 600 nm (OD_{600}) of ~ 0.1 and grown at 30°C. For galactose titration experiments, cells were back-diluted into fresh medium containing the specified galactose concentration. Cells were grown for 6 hours (~ OD_{600} of 0.8–1.0) before measuring GFP levels on a Cell Lab Quanta SC flow cytometer (Beckman Coulter, Fullerton, CA).

Fluorescence quantification

Population averaged fluorescence values were measured on a Quanta flow cytometer with the following settings: 488 nm laser line, 525 nm bandpass filter, and photomultiplier tube setting of 7.53 on FL1 (GFP) and 6.53 on FL3 (RFP). Fluorescence data were collected under low flow rates for ~ 20,000 viable cells. Cells bearing plasmids not expressing RFP were used to set a "RFP negative" gate. Viable cells bearing the plasmid were selected by gating for cells with fluorescence values on FL3 greater than the RFP negative gate. GFP fluorescence levels were determined from 10,000 counts in this selected population. Since the pTEF1-yEGFP cassette is integrated into the chromosome, the mean FL1 values from the entire population of viable, RFP positive cells was determined as the sample's GFP expression level. Relative fluorescence levels

were calculated by normalizing mean GFP levels by the mean GFP level of the no ribozyme control. All fluorescence data are reported as the mean ± SD from at least three independent experiments.

Acknowledgements

Special thanks to Kristy Hawkins for providing strain CSY132 and pCS346 and for technical assistance in building and integrating the expression cassette for CSY341. Funding generously provided by the Joseph J. Jacobs Institute of Molecular Engineering for Medicine (Caltech) and the Center for Biological Circuit Design (Caltech).

References

1. Isaacs FJ, Dwyer DJ, Collins JJ: **RNA synthetic biology.** *Nat Biotechnol* 2006, **24:**545 – 554.
2. Kobayashi H, Kaern M, Araki M, Chung K, Gardner TS, Cantor CR, Collins JJ: **Programmable cells: interfacing natural and engineered gene networks.** *Proc Natl Acad Sci U S A* 2004, **101:**8414 – 8419.
3. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC: **Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.** *Nature* 1998, **391:**806 – 811.
4. Li MJ, Kim J, Li S, Zaia J, Yee JK, Anderson J, Akkina R, Rossi JJ: **Long-term inhibition of HIV-1 infection in primary hematopoietic cells by lentiviral vector delivery of a triple combination of anti-HIV shRNA, anti-CCR5 ribozyme, and a nucleolar-localizing TAR decoy.** *Mol Ther* 2005, **12:**900 – 909.
5. Aagaard L, Rossi JJ: **RNAi therapeutics: principles, prospects and challenges.** *Adv Drug Deliv Rev* 2007, **59:**75 – 86.
6. Grimm D, Streetz KL, Jopling CL, Storm TA, Pandey K, Davis CR, Marion P, Salazar F, Kay MA: **Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways.** *Nature* 2006, **441:**537 – 541.
7. Hanahan D, Weinberg RA: **The hallmarks of cancer.** *Cell* 2000, **100:**57 – 70.
8. Qi M, Elion EA: **MAP kinase pathways.** *J Cell Sci* 2005, **118:**3569 – 3572.

9. McCormick F: **Signalling networks that cause cancer.** *Trends Cell Biol* 1999, **9:**M53 – 56.

10. Long DM, Uhlenbeck OC: **Self-cleaving catalytic RNA.** *Faseb J* 1993, **7:**25 – 30.

11. Khvorova A, Lescoute A, Westhof E, Jayasena SD: **Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity.** *Nat Struct Biol* 2003, **10:**708 – 712.

12. Blount KF, Uhlenbeck OC: **The structure-function dilemma of the hammerhead ribozyme.** *Annu Rev Biophys Biomol Struct* 2005, **34:**415 – 440.

13. Salehi-Ashtiani K, Szostak JW: **In vitro evolution suggests multiple origins for the hammerhead ribozyme.** *Nature* 2001, **414:**82 – 84.

14. Hammann C, Norman DG, Lilley DM: **Dissection of the ion-induced folding of the hammerhead ribozyme using 19F NMR.** *Proc Natl Acad Sci U S A* 2001, **98:**5503 – 5508.

15. De la Pena M, Gago S, Flores R: **Peripheral regions of natural hammerhead ribozymes greatly increase their self-cleavage activity.** *Embo J* 2003, **22:**5561 – 5570.

16. Pley HW, Flaherty KM, McKay DB: **Three-dimensional structure of a hammerhead ribozyme.** *Nature* 1994, **372:**68 – 74.

17. Canny MD, Jucker FM, Kellogg E, Khvorova A, Jayasena SD, Pardi A: **Fast cleavage kinetics of a natural hammerhead ribozyme.** *J Am Chem Soc* 2004, **126:**10848-10849.

18. Penedo JC, Wilson TJ, Jayasena SD, Khvorova A, Lilley DM: **Folding of the natural hammerhead ribozyme is enhanced by interaction of auxiliary elements.** *Rna* 2004, **10:**880 – 888.

19. Saksmerprome V, Roychowdhury-Saha M, Jayasena S, Khvorova A, Burke DH: **Artificial tertiary motifs stabilize trans-cleaving hammerhead ribozymes under conditions of submillimolar divalent ions and high temperatures.** *Rna* 2004, **10:**1916 – 1924.

20. Weinberg MS, Rossi JJ: **Comparative single-turnover kinetic analyses of trans-cleaving hammerhead ribozymes with naturally derived non-conserved sequence motifs.** *FEBS Lett* 2005, **579:**1619 – 1624.

21. Weinberg M, Passman M, Kew M, Arbuthnot P: **Hammerhead ribozyme-mediated inhibition of hepatitis B virus X gene expression in cultured cells.** *J Hepatol* 2000, **33:**142 – 151.

22. Weinberg MS, Ely A, Passman M, Mufamadi SM, Arbuthnot P: **Effective anti-hepatitis B virus hammerhead ribozymes derived from multimeric precursors.** *Oligonucleotides* 2007, **17:**104 – 112.

23. Grassi G, Forlino A, Marini JC: **Cleavage of collagen RNA transcripts by hammerhead ribozymes in vitro is mutation-specific and shows competitive binding effects.** *Nucleic Acids Res* 1997, **25:**3451 – 3458.

24. Scherr M, Grez M, Ganser A, Engels JW: **Specific hammerhead ribozyme-mediated cleavage of mutant N-ras mRNA in vitro and ex vivo. Oligoribonucleotides as therapeutic agents.** *J Biol Chem* 1997, **272:**14304 – 14313.

25. Pan WH, Xin P, Morrey JD, Clawson GA: **A self-processing ribozyme cassette: utility against human papillomavirus 11 E6/E7 mRNA and hepatitis B virus.** *Mol Ther* 2004, **9:**596 – 606.

26. Win MN, Smolke CD: **A modular and extensible RNA-based gene-regulatory platform for engineering cellular function.** *Proc Natl Acad Sci U S A* 2007.

27. Michienzi A, Rossi JJ: **Intracellular applications of ribozymes.** *Methods Enzymol* 2001, **341:**581 – 596.

28. Hawkins KM, Smolke CD: **The regulatory roles of the galactose permease and kinase in the induction response of the GAL network in Saccharomyces cerevisiae.** *J Biol Chem* 2006, **281:**13485 – 13492.

29. Alifano P, Bruni CB, Carlomagno MS: **Control of mRNA processing and decay in prokaryotes.** *Genetica* 1994, **94:**157 – 172.

30. Smolke CD, Carrier TA, Keasling JD: **Coordinated, differential expression of two genes through directed mRNA cleavage and stabilization by secondary structures.** *Appl Environ Microbiol* 2000, **66:**5399 – 5405.

31. Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, Khvorova A: **Rational siRNA design for RNA interference.** *Nat Biotechnol* 2004, **22:**326 – 330.

32. Heale BS, Soifer HS, Bowers C, Rossi JJ: **siRNA target site secondary structure predictions using local stable substructures.** *Nucleic Acids Res* 2005, **33:**e30.

33. Scherr M, Rossi JJ: **Rapid determination and quantitation of the accessibility to native RNAs by antisense oligodeoxynucleotides in murine cell extracts.** *Nucleic Acids Res* 1998, **26:**5079 – 5085.

34. Mori N, Kyo S, Nakamura M, Hashimoto M, Maida Y, Mizumoto Y, Takakura M, Ohno S, Kiyono T, Inoue M: **Expression of HER-2 affects patient survival and paclitaxel sensitivity in endometrial cancer.** *Br J Cancer* 2010, **103:**889 – 898.

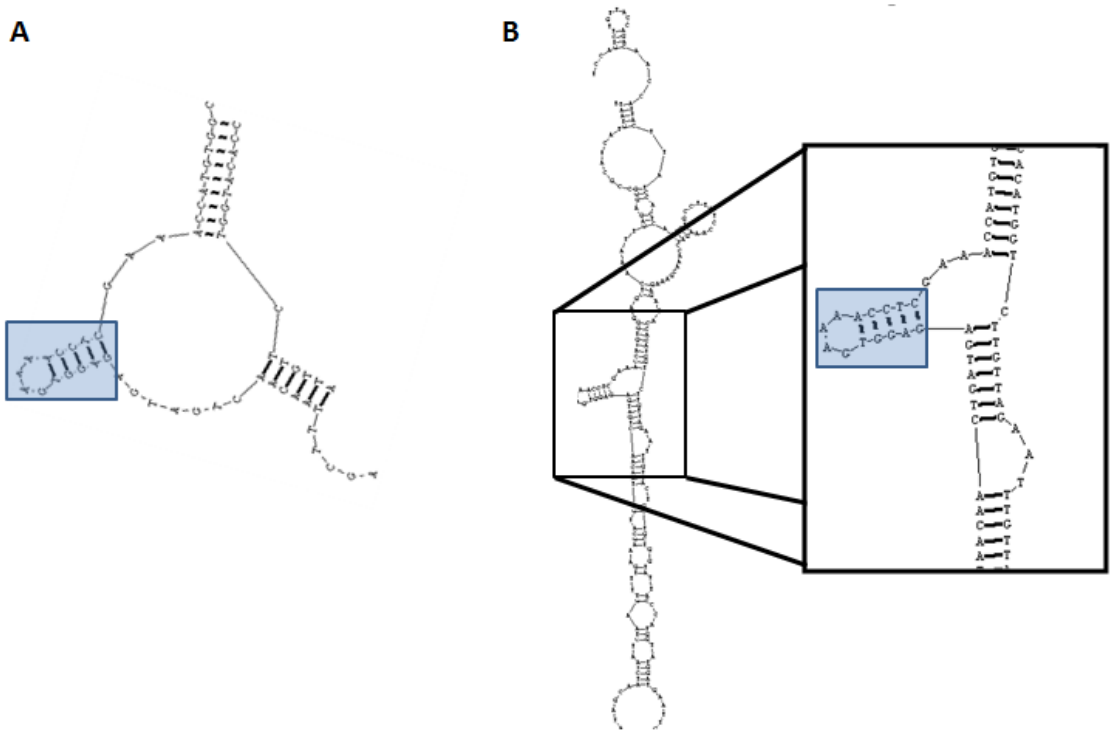
35. Negro A, Brar BK, Lee KF: **Essential roles of Her2/erbB2 in cardiac development and function.** *Recent Prog Horm Res* 2004, **59:**1 – 12.

36. Levine MN: **Trastuzumab cardiac side effects: only time will tell.** *J Clin Oncol* 2005, **23:**7775 – 7776.

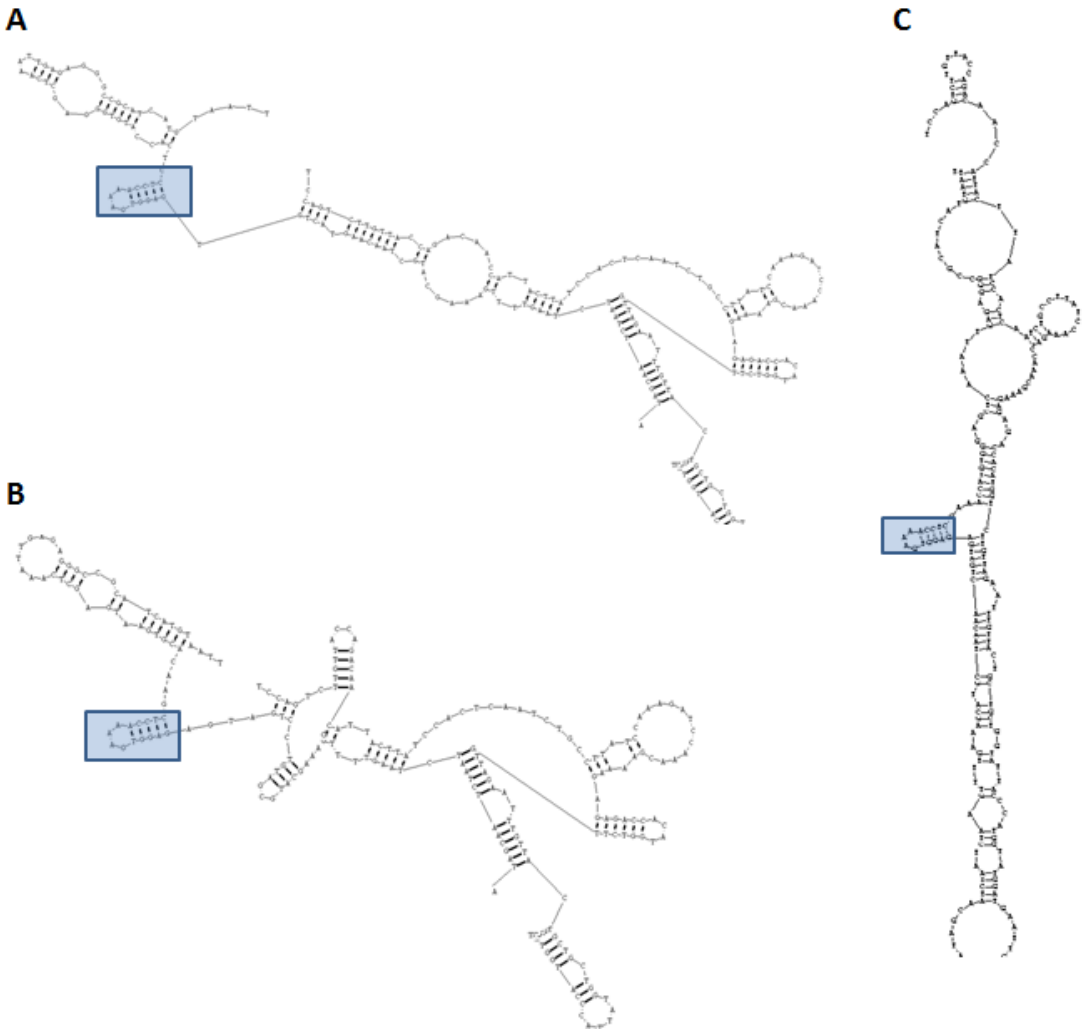
37. Sambrook J RD: *Molecular Cloning: A Laboratory Manual, 3rd edn.* Cold Spring Harbor, NY: Cold Spring Harbor Lab Press; 2001.

38. Guthrie C, Fink G: *Guide to Yeast Genetics and Molecular and Cell Biology.* London: Elsevier, Inc.; 2004.

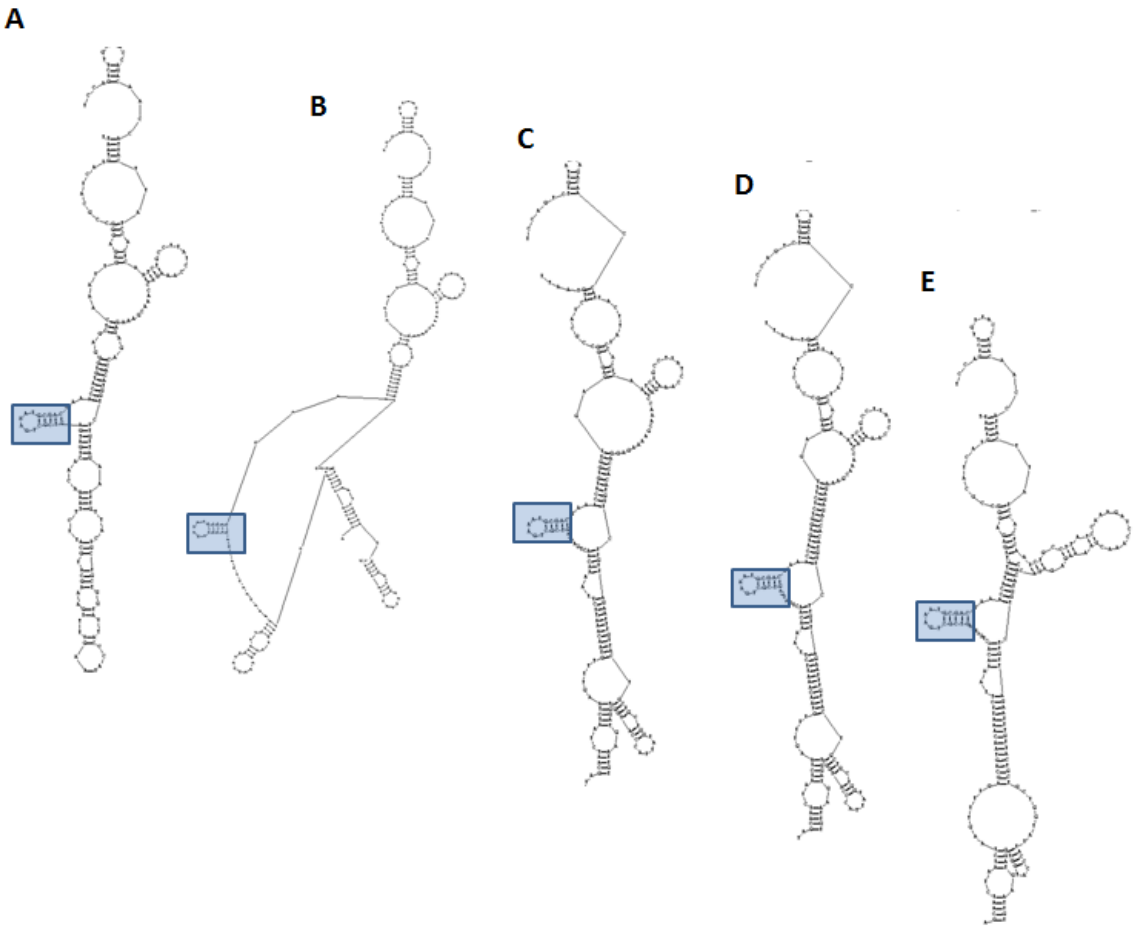
Supplementary figures



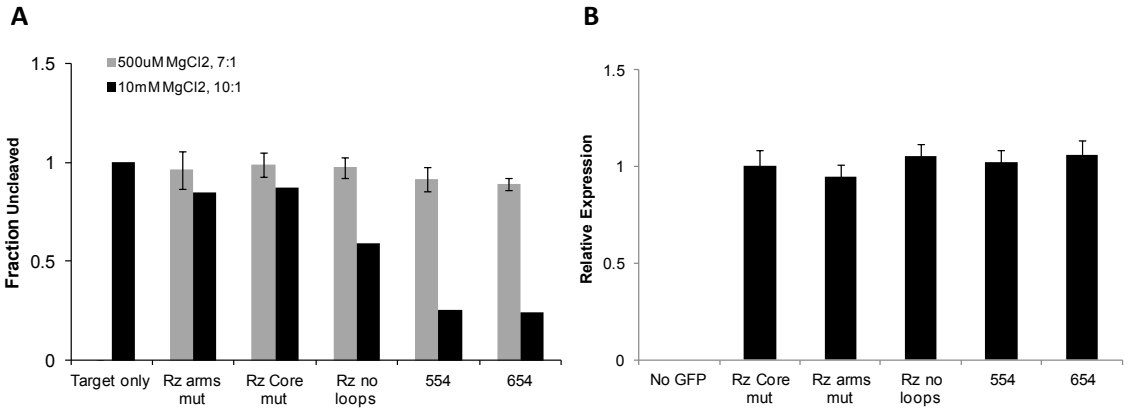
Supplementary Figure 4.1. Comparison of minimal ribozyme folded with truncated and extended target sequence in RNAstructure. A. The minimal ribozyme, SS1, folded without promoter and terminator transcript tails to truncated target sequence shows proper ribozyme-target binding and formation of the catalytic core. **B.** The minimal ribozyme, SS1, folded with promoter and terminator transcript tails from pCS933 to the extended 137 nucleotide (nt) yEGFP sequence shows nucleotides in the catalytic core preferentially binding to target sequence, preventing core formation. Stem II boxed in blue to provide orientation and comparison between figures.



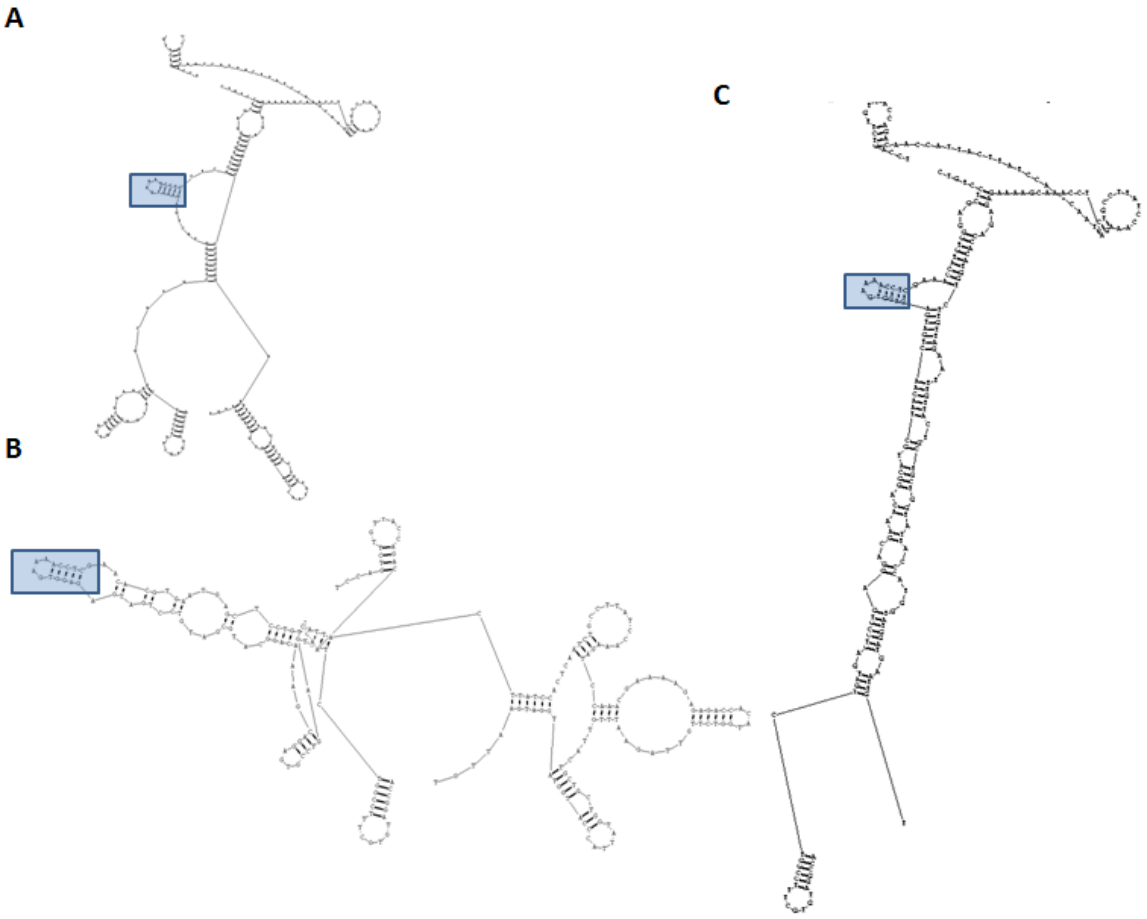
Supplementary Figure 4.2. Control ribozyme sequences from pCS933 folded with extended 137 nucleotide yEGFP target sequence. A. SSCR1, a ribozyme with scrambled core sequence folded with target shows mutated core binding to target sequence. **B.** SSACR, a ribozyme with scrambled targeting arms shows that targeting arms do not properly fold with target. **C.** SS1, a minimal ribozyme without canonical loops, shows nucleotides in the catalytic core binding to target sequence. Stem II boxed in blue to provide orientation and comparison between figures.



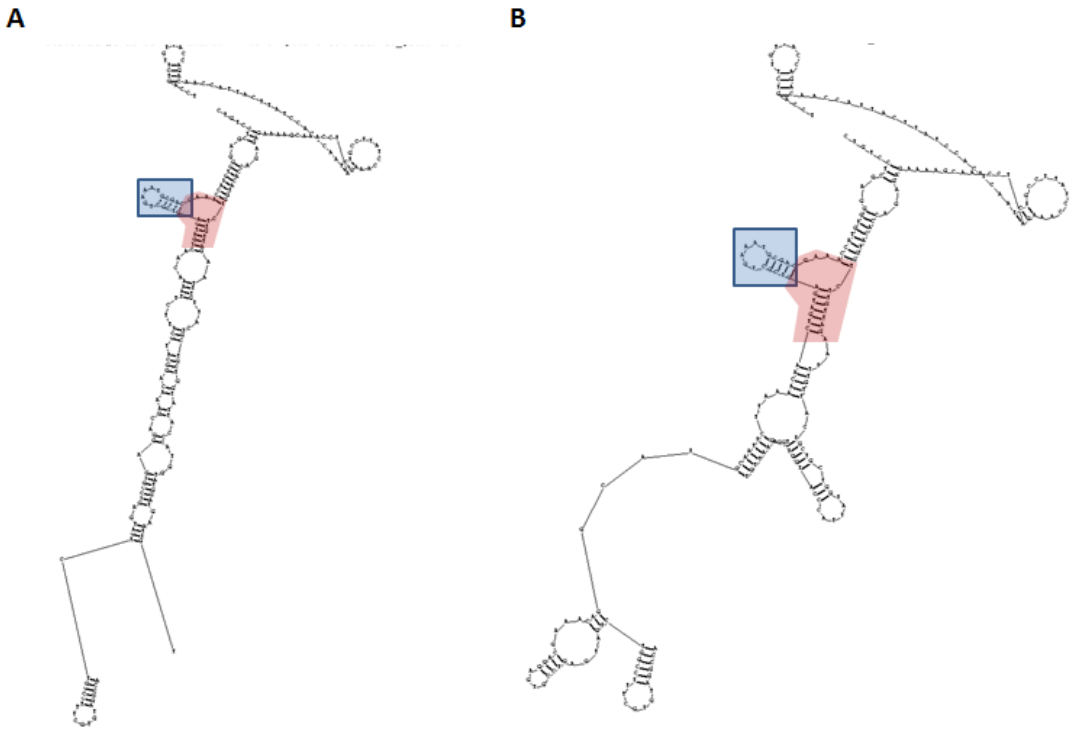
Supplementary Figure 4.3. Canonical ribozyme sequences from pCS933 folded with extended 137 nucleotide yEGFP target sequence. A. 554, a ribozyme with canonical PLMVd stem loops folded with target shows core binding to target sequence. **B.** 654, a ribozyme with canonical PLMVd stem loops folded with target shows 5' targeting arm preferentially binding as a hairpin instead of to target sequence. **C-E.** 1154, 1154+5, and 1654 folded with target show ribozyme binding to target sequence and proper catalytic core formation. Stem II boxed in blue to provide orientation and comparison between figures.



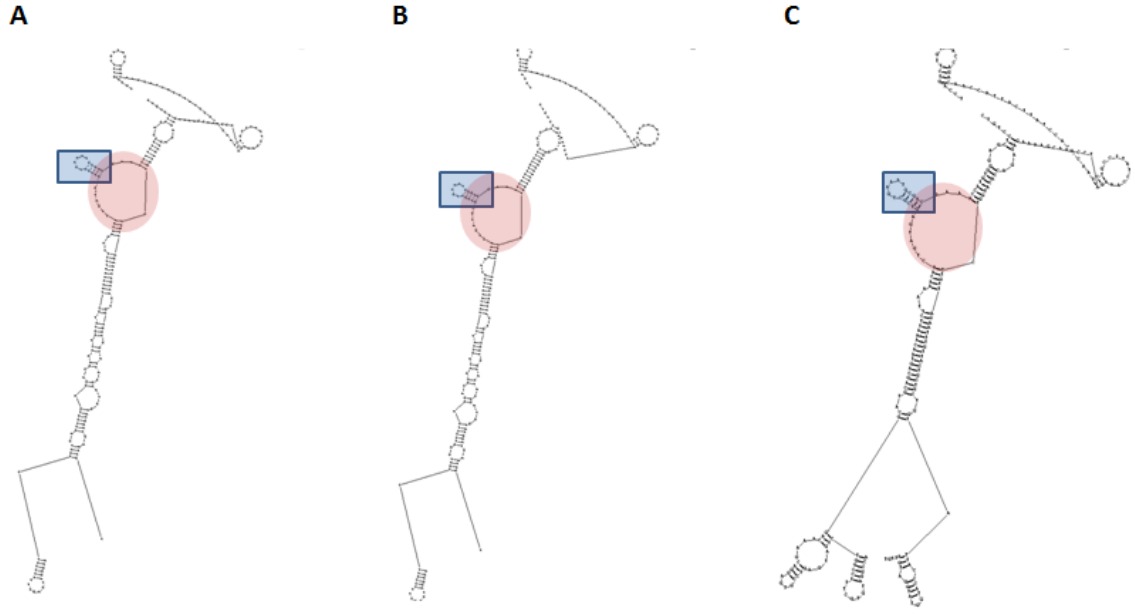
Supplementary Figure 4.4. Initial ribozyme designs with chRZ processing expression cassette (pCS975). A. *In vitro* cleavage assays after 1 hr at various Mg^{+2} concentrations demonstrate the chRz-processing cassette thRzs are highly sensitive to Mg^{+2} concentrations. **B.** *In vivo* fluorescence levels of yEGFP in various controls and constructs. The initial ribozyme designs are insufficient to knockdown expression *in vivo* as expected from the *in vitro* data. Relative fluorescence levels were calculated by normalizing mean GFP levels of cells by the mean GFP level of the no ribozyme control. All fluorescence data are reported as the mean ± SD from at least three independent experiments.



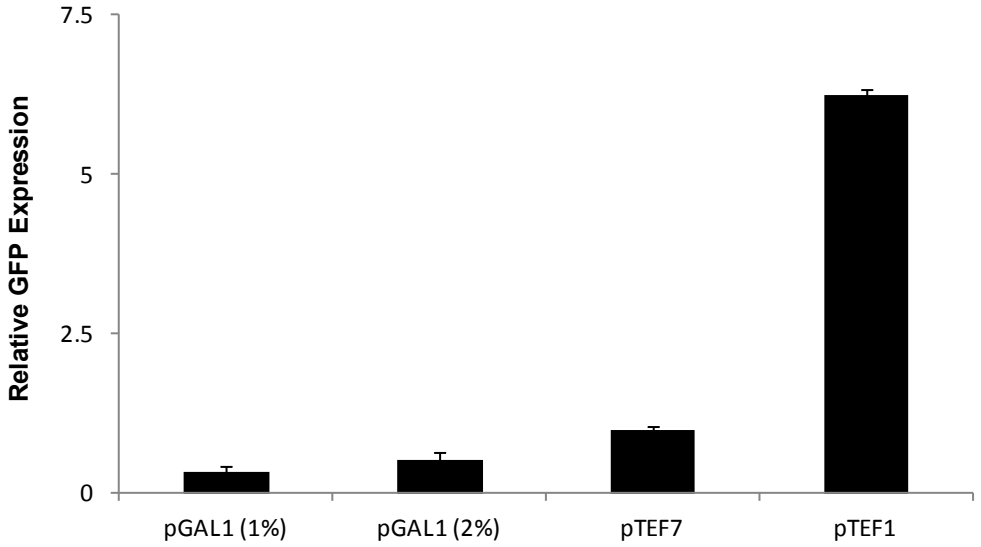
Supplementary Figure 4.5. Control ribozyme sequences from pCS975 folded with extended 137 nucleotide yEGFP target sequence. A. SSCR1, a ribozyme with scrambled core sequence folded with target shows arms correctly binding target sequence. Mutated core also binds "X" in NUX triplet of target sequence. **B.** SSACR, a ribozyme with scrambled targeting arms shows that targeting arms do not properly fold with target. **C.** SS1, a minimal ribozyme without canonical loops, shows nucleotides in the catalytic core binding to target sequence. Stem II boxed in blue to provide orientation and comparison between figures.



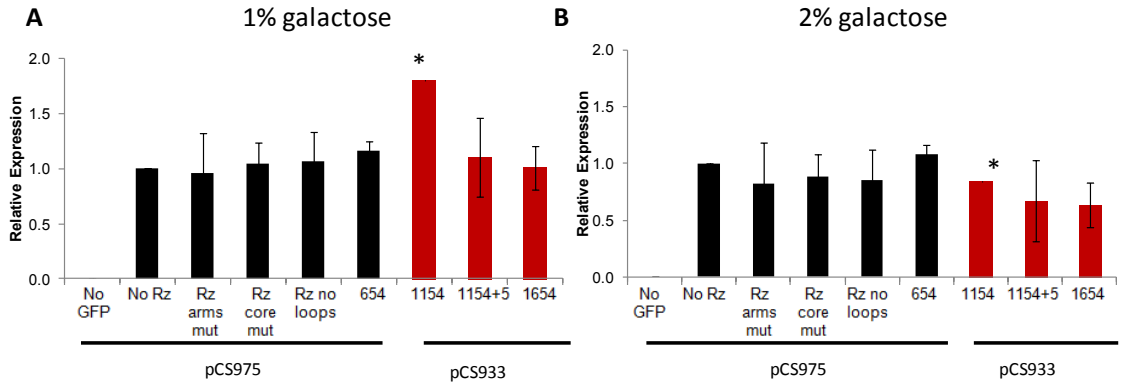
Supplementary Figure 4.6. Canonical ribozyme sequences from pCS975 folded with extended 137 nucleotide yEGFP target sequence. A. 554, a ribozyme with canonical PLMVd stem loops folded with target shows core highlighted in red binding to target sequence. **B.** 654, a ribozyme with canonical PLMVd stem loops folded with target shows core highlighted in red binding to target sequence. Stem II boxed in blue to provide orientation and comparison between figures**.**



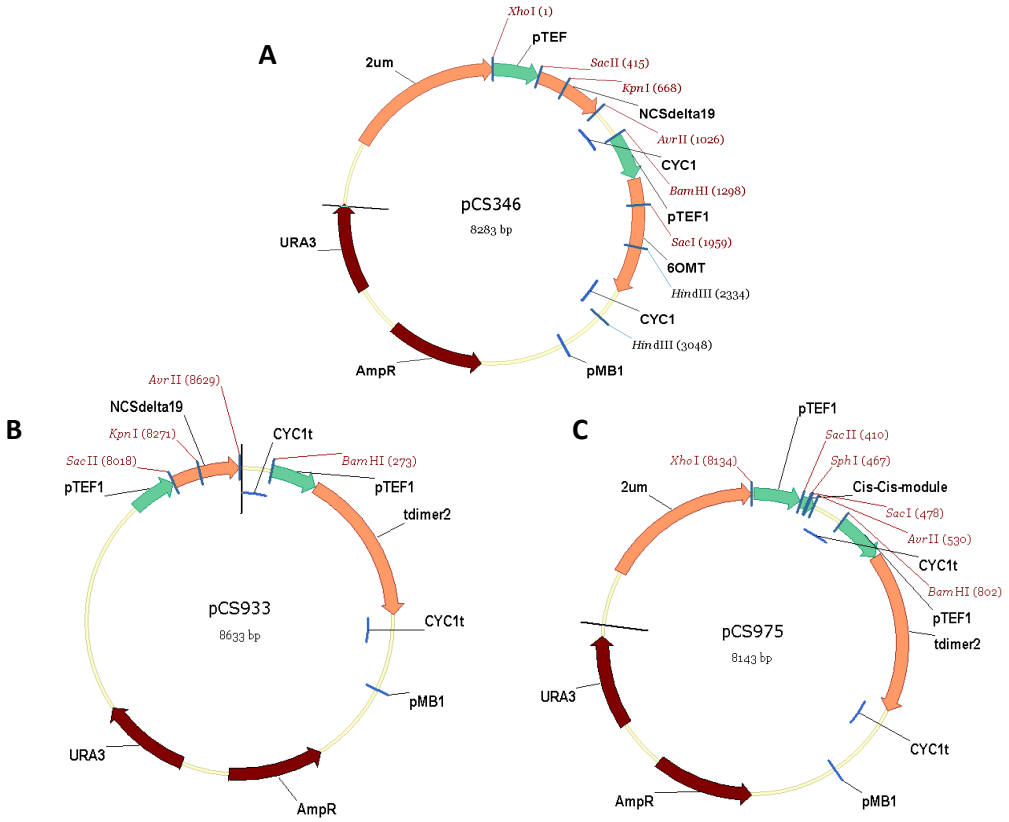
Supplementary Figure 4.7. Canonical ribozyme sequences from pCS975 folded with extended 137 nucleotide yEGFP target sequence. A-C. 1154, 1154+5, and 1654, respectively folded with target show ribozyme binding to target sequence and proper catalytic core formation highlighted in red. Stem II boxed in blue to provide orientation and comparison between figures.



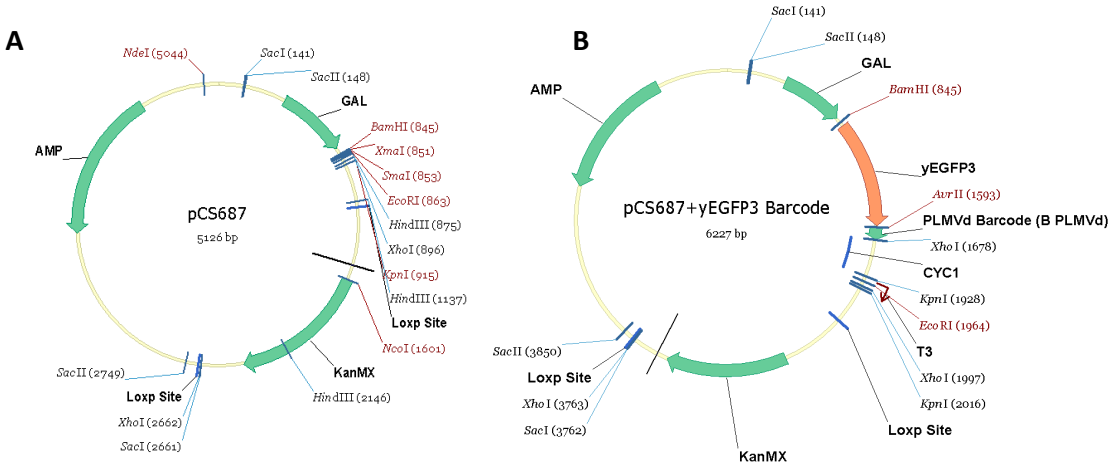
Supplementary Figure 4.8. Promoter characterization. pGAL1 induced at shown percentage of galactose for 6 hours. GFP expression is calculated from the fluorescent geometric mean of viable cells gated above the nonfluorescent control. Expression is shown relative to pTEF7. pTEF1 data from reference for TEF1 library showing pTEF7 is 16% of pTEF1 strength.



Supplementary Figure 4.9. Ribozyme designs with extended arms in pCS933. A. *In vitro* cleavage assays after 1 hr at various Mg^{+2} concentrations. **B.** *In vivo* fluorescence levels of yEGFP in various controls and constructs. Relative fluorescence levels were calculated by normalizing mean GFP levels of cells by the mean GFP level of the no ribozyme control. All fluorescence data are reported as the mean ± SD from at least three independent experiments, except where a single sample is indicated by * and no error bars.



Supplementary Figure 4.10. Plasmid maps for ribozyme expression vectors. A. pCS346 was previously constructed by K. Hawkins with two expression cassettes with promoter (pTEF1) and terminator (CYC1t) controlling two enzymes NCSdelta19 and 6OMT. **B.** pCS933 was constructed from by cloning a version of RFP, tdimer2, into the SalI and NotI restriction sites downstream of the second TEF1 promoter in pCS346 and served as a transformation control signal. **C.** pCS975 was constructed from pCS933 by cloning a cassette containing two chRz with two intervening unique restriction sites, SphI and SacI, into AvrII and SacII using the primers Cis-Cis.sTRSV.FWD and Cis-Cis.sTRSV.REV to create the cis-processing cassette shown as Cis-Cis module.



Supplementary Figure 4.11. Plasmid maps for construction of CSY341 expression vectors. A. pCS687. **B.** pCS1340 was constructed from pCS687 adding yEGFP between BamHI and AvrII and B_PLMVd Barcode between AvrII and XhoI.

Supplementary tables

Supplementary Table 4.1. Plasmids

pCS#	Description	Yeast Marker	E.coli Marker
pCS687	LoxP integrating vector, no yeast origin of replication, derived from pCS270, a pUC vector.	KanMX	AMP
pCS1340	pCS687+ yEGFP cloned between BamHI and AvrII and B_PLMVD (5'-) cloned in 3'UTR between AvrII and XhoI	KanMX	AMP
pCS346	Previously constructed by K. Hawkins with two expression cassettes with promoter (pTEF1) and terminator (CYC1t)	URA	AMP
pCS933	pCS346 + tdimer2 between SalI and NotI	URA	AMP
pCS975	pCS933 + cis-cis.STRSV Cassette between SacII and AvrII	URA	AMP
	pCS933 SSCR1 (aka "Rz core mut")	URA	AMP
	pCS933 SSACR (aka "Rz arms mut")	URA	AMP
	pCS933 SS1-(aka "mRz" or "Rz no loops")	URA	AMP
	pCS933 554	URA	AMP
	pCS933 654	URA	AMP
	pCS933 1154	URA	AMP
	pCS933 1154+5	URA	AMP
	pCS933 1654	URA	AMP
	pCS975 SSCR1 (aka "Rz core mut")	URA	AMP
	pCS975 SSACR (aka "Rz arms mut")	URA	AMP
	pCS975 SS1-(aka "mRz" or "Rz no loops")	URA	AMP
	pCS975 554	URA	AMP
	pCS975 654	URA	AMP
	pCS975 1154	URA	AMP
	pCS975 1154+5	URA	AMP
	pCS975 1654	URA	AMP

Supplementary Table 4.2. Ribozyme sequences

Ribozyme	Type	Sequence	Cloning notes
SSCR1	Rz core mut	GGAAAGCATGCTAACAAGTACTGTGAGGTGAAAACC TCCTCACCATGTGGGAGCTCAAAC	Annealed for insert into pCS933; Annealed and SacI and SphI double-digested for insertion into pCS975
SSACR	Rz arms mut	GGAAAGCATGCGATGTCCTGATGAGAGGTGAAAACC TCGAACACGTGAATGAGCTCAAAC	
SS1 (aka mRz)	Rz no loops	GGAAAGCATGCTAACAACTGATGAGAGGTGAAAACC TCGAAACCATGTGGGAGCTCAAAC	
554	Canonical PLMVd Rz	GGAAAGCATGCATTCTTAAAACAACTGATGAGTCGCT GAAATGCGACGAAACCATGTGGGAGCTCAAAC	
654	Canonical PLMVd Rz	GGAAAGCATGCAATTCTTAAAACAACTGATGAGTCGC TGAAATGCGACGAAACCATGTGGGAGCTCAAAC	
1154	Canonical PLMVd Rz	GGTAACAAATTCTTAAAACAACTGATGAGTCGCTGAA ATGCGACGAAACCATGTGGC	Annealed for insert into pCS933
1154+5	Canonical PLMVd Rz	GGTAACAAATTCTTAAAACAACTGATGAGTCGCTGAA ATGCGACGAAACCATGTGGTCTCTC	
1654	Canonical PLMVd Rz	GGAGCAGTAACAAATTCTTAAAACAACTGATGAGTCG CTGAAATGCGACGAAACCATGTGGC	
1154	Canonical PLMVd Rz	CTAACAAATTCTTAAAACAACTGATGAGTCGCTGAAA TGCGACGAAACCATGTGGGAGCT	Annealed for insert into pCS975
1154+5	Canonical PLMVd Rz	CTAACAAATTCTTAAAACAACTGATGAGTCGCTGAAA TGCGACGAAACCATGTGGTCTCTGAGCT	
1654	Canonical PLMVd Rz	CAGCAGTAACAAATTCTTAAAACAACTGATGAGTCGC TGAAATGCGACGAAACCATGTGGGAGCT	

Supplementary Table 4.3. Yeast strains

CSY#	Description
132	*MATa his3-11 ,15 trp1-1 leu2-3 ura3-1 ade2-1, pTEF1-yEGFP-PEST*
341	*MATa his3-11 ,15 trp1-1 leu2-3 ura3-1 ade2-1,* Δgal2, *pGAL1-yEGFP*

Supplementary Table 4.4. Cloning primers

Name	Sequence	Notes
tdimer2.K2.fwd	AAAAAAGTCGACATTAAATAATGGTGGCCTCCTCCG AGGA	SalI, NotI restriction sites; PCR of tdimer2 for insert into pCS346
tdimer2.rev	AAAAAAGCGGCCGCGAAATTCGCTTATTTAGAAGT	
Cis-Cis. sTRSV.FWD	GGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCC GTGAGGACGAAACAGGCATGCAAAAAGAGCTCCTGT CACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGG ACGAAACAGC	AvrII and SacII restriction sites; Annealed for insert into pCS933
Cis-Cis. sTRSV.REV	CTAGGCTGTTTCGTCCTCACGGACTCATCAGACCGGA AAGCACATCCGGTGACAGGAGCTCTTTTTGCATGCCT GTTTCGTCCTCACGGACTCATCAGACCGGAAAGCAC ATCCGGTGACAGCCGC	
tdimer2.SacII.mut.Fwd	CTACAAGGTGAAGTTCAGAGGCACCAACTTCCCCC	Primers to mutate Sac II sites in tdimer2
tdimer2.SacII.mut.Rev	GGGGGAAGTTGGTGCCTCTGAACTTCACCTTGTAG	
GAL2ko.fwd	ATGCACCTTATTCAATTATCATCAAGAATAGTAATAG TTAAGTAAACACAAGATTAACATAATAGTGCGGGCCT CTTCGCTATTACGCCA	Primers amplifying pCS687 for integration of pGAL1-yEGFP into GAL2 locus
GAL2ko.rev:	ATGATAATTAAAATGAAGAAAAAACGTCAGTCATGA AAAATTAAGAGAGATGATGGAGCGTCTCACTTCACTA TAGGGAGACCGGCAGAT	

Supplementary Table 4.5. *In vitro* assay primers

Name	Sequence	Notes
T7Rbz 5'-23.FWD	**GAAATTAATACGACTCACTATA**GGCATAGCAATCTA ATCTAAGTTTTGC	T7 sequence in bold; 137nt target transcript yEGFP (573-708) 86bp 5' and 51bp 3'
T7Rbz 3'.REV	AATTACATGATGCGGCCCTCTCAACCTA	
Target.yEGFP.FWD	**GAAATTAATACGACTCACTATA**TCCAGTCTTGTTAC CAGACAAC	
Target.yEGFP.REV	CAATTCATCCATACCATGGGTAATACC	
137nt target sequence	TCCAGTCTTGTTACCAGACAACCATTACTTATCCACT CAATCTGCCTTATCCAAAGATCCAAACGAAAAGAGA GACCACATGGTCTTGTTAGAATTTGTTACTGCTGCTG GTATTACCCATGGTATGGATGAATTGT	Actual sequence transcribed

Chapter 5

Conclusion

Robustly engineering cell fate via synthetic circuits offers the potential to direct developmental programs and intervene in aberrantly activated cell processes. Controlling cell fate will advance regenerative medicine by facilitating the *ex vivo* construction of replacement tissues. Further, synthetic circuits may be configured to recognize hyperactive programs that lead to disease and reinstitute the proper behavior or eliminate misbehaving cells on a cell-by-cell basis limiting side-effects of such therapies [1, 2]. However, the engineering of systems to regulate cell fate requires precise control over circuit performance as well as the ability to interface with decision-making pathways [3]. To facilitate the construction of synthetic circuits that regulate cellular behavior for a wide range of potential applications, we developed RNA-based controllers and systems that were applied to regulate signaling in a model MAPK pathway and direct cell fate.

Utilizing modular and tunable RNA-based controllers and network architectures, we constructed networks of RNA-based control systems that interface with the *Saccharomyces cerevisiae* mating pathway to dictate entry into one of three alternative fates dependent on environmental stimuli. In building these networks, we developed a readily translatable method for identifying control points within natural networks that enable the construction of a modular interface between synthetic circuitry and native networks. Additionally, we demonstrated the rational tuning of circuit performance via the exchange of well-defined parts to compose networks capable of actuating changes in cellular behavior in response to environmental cues. Further, we constructed network architectures which facilitate reduced interference from simultaneously integrated opposing programs and identified sensitive parameters for engineering robust circuit performance. This work provides a model for engineering systems that regulate signaling

and direct cells fate which may be applied to additional decision-making pathways to advance tissue engineering strategies, treat diseases, and study the behavior of natural regulatory networks.

Immediate challenges for the development of RNA-based control systems that regulate signaling and dictate cell fate

The constructed molecular network diverters robustly program fate in response to small-molecule input when integrated into cells independent of the opposing diverter. However, due to diverter antagonization, dual-fate routing is only permitted at elevated levels of small-molecule input supported by metabolic modulation.. Diverter antagonization occurs through subthreshold basal levels of expression from the opposing diverter reducing the actuated change in pathway activity induced by the triggered diverter, leading to weaker routing. There exist several potential avenues by which to reduce the basal expression levels from the diverters and minimize diverter antagonization. Reducing the strength of expression by exchanging the promoters that regulate the most antagonistic expression modules will reduce basal expression levels. However, this strategy necessarily reduces the expression induced from the module, potentially hindering diverter performance. Reducing module expression by modification of the plasmid system or integration of the expression modules offer the same limitations as promoter exchange. Nevertheless, integration of the expression modules may provide an advantage over promoter exchange, as integration has been demonstrated to reduce the variance of gene expression compared to plasmid-based systems [4]. While broadly reducing expression from the various modules may reduce the triggered levels of expression from the diverters, the reduction in diverter antagonism may more than

compensate for the reductions, potentially enhancing dual-fate routing performance in these systems.

Secondary layers of control provide an alternative to broadly reducing expression from the diverters. Trans-acting RNA elements have been demonstrated to reduce basal level leakage from circuits in the OFF state, enhancing circuit performance [5-7]. While these particular elements are not amenable to the regulation gene expression in yeast, trans-acting ribozymes offer a control tool for limiting the basal levels of expression in our system. We developed trans-acting ribozymes that target a sequence in the GFP transcript. Appending this GFP target sequence to the genes of pathway regulators may reduce basal expression when trans-ribozymes are coexpressed with the diverters. The development of additional trans-acting ribozymes and cognate target sequences could allow independent regulation of various genes. However, our results indicate that the developed trans-acting ribozymes efficiently regulate expression within a relatively narrow window of transcript expression, presenting limitations in the effective incorporation of these elements to our system. Further work is required to determine if the limited window of efficacy is specific to trans-acting ribozymes in general or to the particular target sequence to which the ribozyme was developed. Additionally, our results suggest that one of the key limitations to trans-ribozyme-mediated knockdown may be facilitating the intermolecular binding event between the trans-acting ribozymes and the target transcript. Co-expression of trans-ribozymes with target from the same locus or vector may overcome these limitations and improve the range of expression levels over which knockdown is observed by facilitating transcript colocalization.

Given the success of their cis-acting counter parts as switches, the development of more potent trans-acting ribozymes would poise these elements as expression platforms by which to construct ligand-responsive trans-acting RNA-based controllers. Trans-acting RNA-based controllers that regulate gene expression in response to small molecule concentrations have been demonstrated in mammalian systems [8, 9] With these trans-acting switches, basal expression levels could be regulated in response to small-molecule concentrations by appending the target sequence to the desired genes. Conditionally reducing target expression via the presence or absence of the small molecule offers a mechanism for reducing basal expression levels from the various modules without a similar reduction in the triggered levels. Reduced diverter antagonism and greater fold changes in expression from the diverters may be achieved by the development and incorporation of ligand-responsive trans-acting RNA controllers.

In addition to layers of trans-acting controllers, the introduction of mutual inhibitory modules between the dual diverters may facilitate enhanced dual-routing performance by reducing diverter antagonization. Mutually inhibitory loops have been demonstrated to allow cells to toggle between two distinct phenotypes [10, 11] and may be constructed from transcriptional or post-transcriptional elements. Construction of transcriptional mutually inhibitory modules requires two orthogonal repressible promoters each regulating the expression of the other's repressor protein. Transcriptional mutual inhibitory modules are simply connected to the diverters by promoter exchange in the expression modules containing constitutive promoters. While the system is readily constructed, identifying repressible promoters of the requisite strength and tuning the mutual inhibitory module present a critical hurdle for enhancing diverter function via

transcriptional mutually inhibitory modules. Construction of repressible promoter libraries will increase the feasibility of constructing and connecting transcriptional mutual inhibitory modules to synthetic circuits. Alternatively, post-transcriptional controllers provide a promoter-independent mechanism for introducing mutual inhibitory loops into the diverter networks, providing greater design flexibility. Post-transcriptional mutually inhibitory loops may be constructed by directly modifying the constructed expression modules with target sites specific for a trans-acting RNA regulatory element. Imbedding cognate trans-acting RNA regulators in the expression modules of the opposing diverter establishes an inhibitory regulatory loop from one diverter to another. Addition of a secondary inhibitory loop with a different trans-acting RNA regulator and cognate target site results in a mutually inhibitory architecture. Further, mutual inhibition via trans-acting RNA regulators can be made ligand-dependent by the use of small molecule-responsive trans-acting RNA switches. The continued development of ligand-responsive switches from potent trans-acting ribozymes, microRNAs, and other trans-acting RNA-based regulators of gene expression will enable the construction of more complex control schemes within synthetic networks. While regulation via post-transcriptional control may be limited in stringency compared to transcriptional control, the added flexibility, tunability, and potential for ligand-dependent regulation position post-transcriptional mutually inhibitory loops to be important tools for the construction of synthetic circuits. The addition of mutually inhibitory modules that enhance differentiation between synthetic circuits' ON and OFF state outputs will enable robust control of cell fate, expanding the range of systems and applications to which synthetic circuits may be applied.

Future directions for the development of RNA-based control systems to regulate signaling and dictate cell fate

The *ex vivo* construction of tissues, the implementation of autonomous immune surveillance, and the development of 'smart' therapeutics will require flexible, tunable, and scalable synthetic networks that can interface with a range of native pathway via circuitry capable of extracting cellular information and actuating changes in cellular behavior. Our work presents a method for identifying control points in native pathways that is amenable across a wide array of systems. Further, the development of the molecular network diverters provides a framework for constructing and tuning synthetic networks that control cellular behavior that is readily transportable across pathways, organisms, and applications. Yet there remain several challenges to extending these systems to applications in health and medicine.

Application of the molecular diverter strategy for biomedical purposes will require the identification of titratable pathway regulators in the decision-making pathways in mammalian cells. While our results with the yeast MAPK pathways illuminate potential candidates within the homologous mammalian pathways, titration studies must be performed in these systems to validate potential pathway regulators' ability to route cells to alternative fates. Further, examination of the pathway response curve to varying regulator expression will facilitate identification of the transitory range over which fate diverges and enable rational design of molecule network diverters. Titration studies to identify pathway regulators may be performed in mammalian cells using the well-documented tetracycline-inducible promoter [12]. Already similar work has been performed using the tetracycline-inducible system to facilitate the programming

of induced pluripotent stem cells (iPSCs) via the control of key transcription factors [13]. In addition to the requisite pathway regulators, our work highlights the utility of pathway-responsive promoters for the construction of feedback loops that amplify and attenuate pathway activity. Identification of relevant pathway-responsive promoters will facilitate the construction of molecular network diverters within new pathways. In the ERK signaling pathway which is homologous to the yeast mating pathway and responsible for controlling cell growth, several pathway-responsive promoters have been characterized and reporter constructs with these elements and a host of other signaling pathway are commercially available [14]. Tuning promoter expression to achieve the desired strength may be accomplished by the construction of promoter libraries or by rational design [15, 16].

Beyond pathway regulators and promoters, improvements in the stringency and regulatory range of RNA-based switches as well as the development of new switch sensor domains will expand the applicability of molecular network diverters. The incorporation of RNA switches exhibiting low basal levels and large dynamic ranges may increase between the range of circuit output, enhancing differentiation between triggered and non-triggered cells. However, computational models of RNA switches have indicated a tradeoff in tuning the stringency of a switch and its sensitivity to the input ligand, ultimately impacting the dynamic range [17]. Nevertheless, selection of switches with high stringency and large dynamic ranges is important particularly for the development of switches responsive to new ligands, an important step in translating these systems to clinical applications. In this study, we incorporated switches responsive to theophylline and tetracycline in our designs. While these molecules are suitable for the regulation of

yeast gene expression, cytotoxicity restricts the concentrations that may be safely administered in mammalian systems thus reducing the range of expression from these elements in a therapeutically relevant context [2]. The selection of new aptamer sequences to FDA-approved molecules will facilitate the construction of switches that may be applied as therapies. Additionally, the development of new aptamer sequences to cellular molecules of interest offers the potential to connect information from endogenous networks to synthetic control systems. The construction of circuits responsive to endogenous molecules will facilitate autonomous cell-based therapies. Synthetic RNA aptamers have been generated *de novo* to various small-molecule and protein targets through *in vitro* selection or SELEX strategies [18, 19]. The Smolke laboratory has generated RNA aptamers that exhibit varying specificities to benzylisoquinoline alkaloids [20] and folinic acid derivates, and is developing high-throughput strategies for the direct selection and characterization of new protein- and small-molecule-responsive aptamers. The development of switches responsive to new molecules will provide additional independent channels of regulation by which to construct orthogonal control of multiple genes within complex networks.

Finally, connecting additional synthetic circuitry to the molecular network diverters may facilitate more complex and sophisticated cellular programming [3]. Complex functions such as the synchronization of genetic clocks and edge detection have been demonstrated via the rational coupling of communication circuits to oscillators and light-responsive logical circuits, respectively, [21, 22]. Coupling synthetic circuits that control cellular signaling to communication circuits may facilitate robust spatial patterning of cell fate, enable tissue homeostasis, and allow the coordination of distributed tasks, requisite achievements for the construction, maintenance, and function

of complex tissues [23-25]. Further, advances in iPSC technology may enable the construction of immunologically compatible tissues from cells synthetically programmed for precise spatial and temporal differentiation on designer cell scaffolds [26-28].

In this work we have demonstrated the construction of RNA-based control systems that route cells to divergent cell fates in response to exogenously applied triggers. These systems highlight the potential to develop synthetic networks that spatially and temporally program cell fate advancing the fields of tissue engineering and molecular medicine. As the systems biology and synthetic biology tool box expands, the future development of these systems promises the construction of larger, more complex networks engineered to control a range of systems from cells to organs to whole organisms for a wide array of biotechnological and medical applications.

References

1. Culler SJ, Hoff KG, Smolke CD: **Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins.** *Science* 2010, **330:**1251 – 1255.
2. Chen YY, Jensen MC, Smolke CD: **Genetic control of mammalian T-cell proliferation with synthetic RNA regulatory systems.** *Proc Natl Acad Sci U S A* 2010, **107:**8531 – 8536.
3. Chen YY, Galloway KE, Smolke CD: **Synthetic biology: advancing biological frontiers by building synthetic systems.** *Genome Biol* 2012, **13:**240.
4. Michener JK, Thodey K, Liang JC, Smolke CD: **Applications of genetically-encoded biosensors for the construction and control of biosynthetic pathways.** *Metab Eng* 2012, **14:**212 – 222.
5. Deans TL, Cantor CR, Collins JJ: **A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells.** *Cell* 2007, **130:**363 – 372.
6. Xie Z, Wroblewska L, Prochazka L, Weiss R, Benenson Y: **Multi-input RNAi-based logic circuit for identification of specific cancer cells.** *Science* 2011, **333:**1307 – 1311.
7. Callura JM, Dwyer DJ, Isaacs FJ, Cantor CR, Collins JJ: **Tracking, tuning, and terminating microbial physiology using synthetic riboregulators.** *Proc Natl Acad Sci U S A* 2010, **107:**15898 – 15903.

8. Beisel CL, Chen YY, Culler SJ, Hoff KG, Smolke CD: **Design of small molecule-responsive microRNAs based on structural requirements for Drosha processing.** *Nucleic Acids Res* 2011, **39:**2981 – 2994.

9. Beisel CL, Bayer TS, Hoff KG, Smolke CD: **Model-guided design of ligand-regulated RNAi for programmable control of gene expression.** *Mol Syst Biol* 2008, **4:**224.

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

11. Kramer BP, Viretta AU, Daoud-El-Baba M, Aubel D, Weber W, Fussenegger M: **An engineered epigenetic transgene switch in mammalian cells.** *Nat Biotechnol* 2004, **22:**867 – 870.

12. Stieger K, Belbellaa B, Le Guiner C, Moullier P, Rolling F: **In vivo gene regulation using tetracycline-regulatable systems.** *Adv Drug Deliv Rev* 2009, **61:**527 – 541.

13. Kozlova EN, Berens C: **Guiding differentiation of stem cells in vivo by tetracycline-controlled expression of key transcription factors.** *Cell Transplant* 2012.

14. Wong KK: **Recent developments in anti-cancer agents targeting the Ras/Raf/ MEK/ERK pathway.** *Recent Pat Anticancer Drug Discov* 2009, **4:**28 – 35.

15. Alper H, Fischer C, Nevoigt E, Stephanopoulos G: **Tuning genetic control through**

promoter engineering. *Proc Natl Acad Sci U S A* 2005, **102:**12678 – 12683.

16. Blount BA, Weenink T, Vasylechko S, Ellis T: **Rational diversification of a promoter providing fine-tuned expression and orthogonal regulation for synthetic biology.** *PLoS One* 2012, **7:**e33279.

17. Beisel CL, Smolke CD: **Design principles for riboswitch function.** *PLoS Comput Biol* 2009, **5:**e1000363.

18. Tuerk C, Gold L: **Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase.** *Science* 1990, **249:**505 – 510.

19. Bowser MT: **SELEX: just another separation?** *Analyst* 2005, **130:**128 – 130.

20. Win MN, Klein JS, Smolke CD: **Codeine-binding RNA aptamers and rapid determination of their binding constants using a direct coupling surface plasmon resonance assay.** *Nucleic Acids Res* 2006, **34:**5670 – 5682.

21. Danino T, Mondragon-Palomino O, Tsimring L, Hasty J: **A synchronized quorum of genetic clocks.** *Nature* 2010, **463:**326 – 330.

22. Tabor JJ, Salis HM, Simpson ZB, Chevalier AA, Levskaya A, Marcotte EM, Voigt CA, Ellington AD: **A synthetic genetic edge detection program.** *Cell* 2009, **137:**1272 – 1281.

23. Weber W, Daoud-El Baba M, Fussenegger M: **Synthetic ecosystems based on airborne inter- and intrakingdom communication.** *Proc Natl Acad Sci U S A* 2007, **104:**10435 – 10440.

24. Weber W, Fussenegger M: **Design of synthetic mammalian quorum-sensing systems.** *Methods Mol Biol* 2011, **692:**235 – 249.

25. Weber W, Schuetz M, Denervaud N, Fussenegger M: **A synthetic metabolite-based mammalian inter-cell signaling system.** *Mol Biosyst* 2009, **5:**757 – 763.

26. Maherali N, Ahfeldt T, Rigamonti A, Utikal J, Cowan C, Hochedlinger K: **A high-efficiency system for the generation and study of human induced pluripotent stem cells.** *Cell Stem Cell* 2008, **3:**340 – 345.

27. Kueh HY, Rothenberg EV: **Regulatory gene network circuits underlying T cell development from multipotent progenitors.** *Wiley Interdiscip Rev Syst Biol Med* 2011.

28. Thomson M, Liu SJ, Zou LN, Smith Z, Meissner A, Ramanathan S: **Pluripotency factors in embryonic stem cells regulate differentiation into germ layers.** *Cell* 2011, **145:**875 – 889.