# Synthetic Gene Circuits in Intestinal Bacteria: Biosensors and Pulse Counters

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:42029544">http://nrs.harvard.edu/urn-3:HUL.InstRepos:42029544</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Synthetic Gene Circuits in Intestinal Bacteria:
Biosensors and Pulse Counters

A dissertation presented
by
Alexander Dimitri Naydich
to
The Harvard John A. Paulson School of Engineering and Applied Sciences

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Engineering Sciences

Harvard University
Cambridge, Massachusetts

April 2019
Abstract

The human microbiome comprises a complex community of microorganisms, residing primarily in the intestine. Variations in the composition of the microbiome have been linked to an increasingly wide array of diseases, leading to a growing interest in the development of microbiome-based clinical interventions. One promising opportunity is the engineering of commensal bacteria to create new strains that can sense disease in the gut and produce therapeutic compounds in response. This dissertation discusses motivations for the construction of synthetic biological circuits in intestinal bacteria and presents developments in two key areas. Because the design of disease-responsive circuits is limited by a relatively small pool of known biosensors, there is a need for expanding the capacity of engineered bacteria to sense and respond to the host environment. Chapter 2 describes the application of a robust genetic memory circuit to identify new bacterial biosensors responding in the healthy and diseased mammalian gut, which may be used to construct diagnostic or therapeutic circuits. Chapter 3 presents novel synthetic circuits that can record multiple pulses of a specific signal, a fundamental capability with many potential applications in biosensing, which has not been reliably demonstrated in synthetic biology thus far. While the clinical application of engineered bacterial strains is still an emerging concept, the advances described herein expand the capabilities of synthetic biology for the creation of increasingly complex, environmentally-responsive systems, representing a practical step toward the construction of clinically useful synthetic tools.
## Contents

Abstract iii

Acknowledgements vi

### Chapter 1: Introduction 1

Preface 2

1.1 From recombinant DNA to synthetic biology 2

1.2 Engineering the microbiome 5

1.3 Bacterial sensors for the intestinal environment 9

1.4 Circuits for processing & recording signals 20

1.5 Aims of this study 23

### Chapter 2: A synthetic memory circuit enables systems-level biosensor discovery at the host–microbe interface 24

Preface 25

2.1 Abstract 25

2.2 Introduction 26

2.3 Results 28

2.4 Discussion 46

2.5 Materials and Methods 50
Chapter 3: Falling-edge pulse counters record repeated discrete occurrences of a stimulus 56

Preface 57

3.1 Abstract 57

3.2 Introduction 58

3.3 Results 61

3.4 Discussion 84

3.5 Materials and methods 87

Chapter 4: Conclusion 92

4.1 In vivo use of engineered bacteria 93

4.2 Broader considerations for recombinant DNA technology 96

Appendix A: Supplemental Material for Chapter 2 100

Appendix B: Supplemental Material for Chapter 3 106

Bibliography 110
Acknowledgements

I have been exceedingly fortunate to have Pam Silver as my research advisor in the course of my graduate work. In addition to allowing me the freedom to take on a variety of fascinating projects, Pam welcomed me into a group in which I found a wealth of experience to build upon, wonderful colleagues of diverse scientific and social backgrounds, and plenty of opportunities for collaboration. I would also like to thank my committee members, Donald Ingber and Neel Joshi, and I am especially grateful to Neel for introducing me to the world of synthetic biology. I would also like to acknowledge the NSF Graduate Research Fellowship Program, which made my work possible.

In the Silver lab, I cannot overstate my gratitude to David Riglar for being an absolutely irreplaceable mentor, advocate and research partner. I would also like to thank Shannon Nangle and Finn Stirling, who were fantastic collaborators with whom I was able to significantly expand the scope of my work. Thank you also to Bryan Hsu and Isaac Plant for always being there to advise, commiserate or celebrate as appropriate. Likewise, my baymates Suhyun Kim, Wiggert Altenburg and Michael Melfi made Bay 5 an entertaining place to be, due in part to their apropos utilization of the Sound Machine. I am also proud to call all of the above people my friends, and I am further grateful to everyone in the Silver lab for the coffees, lunches, happy hours, board game nights, and ski trips that make it such a special place.

Thank you also to my friends outside the lab—Maxim Belomestnykh, Roman Goloborodko, Li Guo, Andrew Hirschl, Shane Mc Morrow, Ryan Musa, Jeffery Roshko and Lucas Waye—for keeping in touch and for finding the time to get together, regardless of the distance between us.
Lastly, I would not be who I am without my family: my cousin Anna, my aunt Tatyana, my grandmother Yevgeniya, my grandparents Irina and Eduard (who passed away), and my parents Irina and Dimitri. Thank you all for setting an example for me and for your perpetual encouragement and confidence in my ability to succeed. The extent of my gratitude to my parents, especially, is impossible to put into words. To them I can only add thank you for teaching me the importance of hard work and education, while simultaneously instilling the appreciation that there is so much more to life. Finally, it is hard to imagine going through graduate school—or the past decade—without the love and support of my wife, Heather. The fact that we were able to get through medical school and graduate school while living four hours apart is a testament to the strength of our bond, and I am thrilled to start the next chapter of our life together.
Chapter 1: Introduction
Preface

The purpose of this chapter is to provide an overview of the motivations behind synthetic biology and to illustrate why the human microbiome constitutes a promising vehicle for the deployment of synthetic biological systems. This chapter further discusses engineered bacterial sensors with a focus on sensing the host intestinal environment, as well as information processing and recording circuits in bacteria. Heather Edward generously provided comments on this chapter.

“We must not wait for favors from Nature; our task is to seize them from her.”

I.V. Michurin, *Results of Sixty Years’ Work* (1934)

1.1 From recombinant DNA to synthetic biology

The above quotation from the Russian botanist, Ivan Michurin, was inspired by his work in the breeding of frost tolerant varieties of various fruits. Today, this Soviet maxim may be more frequently cited with irony, invoking technological hubris; nonetheless, it embodies humanity’s boldness and continued determination to better its own condition beyond the limits of what may have previously been considered possible. Although Michurin himself was skeptical of genetics throughout his career—and worse yet, his work was later appropriated as an inspiration for Lysenkoism, one of the most damaging anti-science campaigns of the twentieth century—the belief that humans can and should alter biology to suit our needs is as alive as ever.
The remainder of the twentieth century brought great advances in both our understanding of biology and in ways that it could be manipulated. The discovery that genetic information was stored in DNA, the articulation of the central dogma, and studies of gene regulation all furthered our understanding of how organisms carry out their functions and pass down their traits. Later, the ability to precisely alter genes through the use of restriction endonucleases and recombination—as well as the development of fundamental tools such as molecular cloning, the polymerase chain reaction (PCR) and oligonucleotide synthesis—made it possible to isolate and study the functions of individual genes without having to rely on mutation and selection. At the same time, it also became possible to express genes in non-native hosts, launching the field of genetic engineering. The first genetically engineered (GE) organisms were created in the 1970s, and the first pharmaceutical produced through transgenic means—human insulin from *Escherichia coli*—was approved in 1982. Since that time, a variety of pharmaceutical and industrial compounds have been produced in engineered bacteria, yeast, and mammalian cells (1).

In addition to its use in biomanufacturing, recombinant DNA technology has expanded to modify traits in agricultural crops, animals and even humans. Transgenic crops with the ability to resist herbicides or to produce insecticides entered the market in the 1990s and are now ubiquitous across the United States (2). Engineered salmon that can reach harvestable size nearly twice as quickly as their non-engineered counterparts (3) were deemed safe for human consumption in 2015 by the U.S. Food and Drug Administration (FDA) and cleared for sale in 2019 (4). To date, most applications of gene and cell therapy in humans have targeted cancer and inherited disease, with the first gene therapy approved in China in 2003 and the first engineered cell therapy approved by the FDA in 2017 (5).
Synthetic biology is a new field that aims to take the potential of recombinant DNA even further. Rather than adding or modifying individual genes to alter traits in an organism, synthetic biology seeks to create completely new biological systems with programmable logic. The field’s mission is to design and build biological circuits that can allow prokaryotic and eukaryotic cells to sense, process and respond to their environment, with potential applications in biosensing, biomanufacturing and human health (6). Such circuits should consist of characterized parts that can be combined in increasingly complex but still predictable ways (7, 8). In effect, a major goal of synthetic biology is to make biological circuits as modular and scalable as electronic ones.

In the past two decades, several technological advances have accelerated the pace and scale of the design-build-test cycle for biological circuits. These include: the falling cost of DNA sequencing (9) and synthesis (10); the development of new gene editing and assembly methods such as Gibson assembly (11), Golden Gate assembly (12) and CRISPR-Cas (13); and the emergence of automation (14, 15). At the same time the complexity of constructed circuits has increased, enabled by a growing collection of characterized parts, including promoters, terminators, ribosome binding sites, transcription factors, reporters and recombinases. Although the creation of functional circuits still routinely requires the construction and testing of multiple variants, there have been efforts to establish reliable parts libraries (16) and to reduce the effects of genetic and cellular context on a circuit’s function while minimizing burden on the host organism (17–20).

As synthetic biology enters its third decade, its focus has started to move beyond proof-of-concept systems to create circuits that have potential real-world applications and are reliable enough for deployment. While the advances described above can certainly expand the
capabilities of biomanufacturing, one truly novel frontier is the development of programmable, cell-based diagnostics and therapeutics in the form of engineered bacterial and human cells that can be applied in clinical settings (21). Most advances in genetic engineering were first demonstrated in prokaryotes, owing to their reduced complexity and ease of propagation, but many synthetic circuits have since been developed in eukaryotic systems. Furthermore, certain properties of mammalian cells—including their capacity for post-translational modification of proteins and their potentially lower immunogenicity—make them especially promising for clinical applications (22, 23). At the same time, synthetic biology techniques are more developed in prokaryotic systems, and, given the multifaceted role of the human microbiome in health and disease, the engineering of bacteria holds unique clinical potential. This dissertation focuses on the construction of genetic circuits in bacteria with the purpose of interfacing with mammalian systems.

1.2 Engineering the microbiome

The human microbiome is diverse community of commensal and mutualistic microorganisms, consisting mostly of bacteria, living inside the body and on the skin. Recent estimates suggest that the number of bacteria found in a typical human is roughly equal to the number of somatic cells (24), with the majority of these bacteria residing in the intestine. In the past decade, the field of microbiome studies has grown rapidly, spurred by the advent of next-generation sequencing technologies, as well as major initiatives such as the Human Microbiome Project (25) and MetaHIT (26). A typical individual may harbor over 100 bacterial species, though overall taxonomic composition varies greatly, even among healthy individuals (27).
Strikingly, these bacteria have been shown to play a role in many aspects of human health, including immune development, metabolism, and endocrine and neurological function (28). Studies in humans and animal models have revealed distinct shifts in microbiome composition associated with a wide range of diseases, including gastrointestinal, metabolic, immune, cardiovascular, neurologic and other disorders (28). The majority of these observations are correlative, potentially limiting their application as clinical interventions. However, one notable success has been the transfer of a healthy individual’s microbiota through fecal transplantation (FMT) as a remarkably effective way to treat patients with recurrent Clostridium difficile infection (29). More recently, clinical studies have attempted to use FMT to treat inflammatory bowel disease (30), metabolic syndrome (31), and even autism (32), though larger-scale studies are needed to better understand its potential in these other disease areas (33).

Beyond correlative studies of microbiome composition, a growing number of mechanistic studies have elucidated specific pathways through which intestinal bacteria affect the function of the host. For example, short-chain fatty acids produced by the microbiota can regulate host insulin sensitivity by binding to GPR43 receptors in adipose tissue (34), in addition to influencing a wide range of other processes (35). Not all the effects of the microbiota are beneficial. Bacteria in the gut can convert dietary choline and L-carnitine into trimethylamine, which is further oxidized to trimethylamine N-oxide in the liver, which accelerates atherosclerosis (36). Furthermore, the varied effects of the microbiota on the metabolism and absorption of pharmaceutical drugs are of marked clinical interest (37–39).

In addition to studying the natural functions of the microbiome and attempting to alter its composition for clinical benefit, the engineering of new capabilities into commensal microbiota constitutes an increasingly attractive platform for therapeutic and diagnostic
applications. Because of their proximity to the host and their ability to influence many aspects of human health, intestinal bacteria are a promising chassis for the deployment of synthetic biological circuits. Such circuits might enable the construction of diagnostic strains that sense and record the presence of biological and environmental factors in the intestine. These engineered strains could then be analyzed after transit through the gut, providing a non-invasive survey of an obscure environment. Even more compelling is the potential for on-demand production of biological therapies. Bacteria might be engineered with circuits that sense disease-associated signals and produce a localized, therapeutic response in the form of a peptide or protein, which may also reduce any side-effects associated with systemic administration of a compound. Such approaches may be particularly appealing for diseases such as inflammatory bowel disease and colorectal cancer, which are difficult to treat and require costly, invasive procedures for diagnosis.

Figure 1.1 Engineered intestinal bacteria containing synthetic gene circuits may sense, process and respond to one or more physiological or pathological signals in the gut.
One of the first demonstrations of engineering of bacteria for \textit{in vivo} therapeutic production was the use of \textit{Lactococcus lactis} to produce IL-10 in the intestine (40). Since then, there have been many studies which have used engineered bacteria, both inside and outside the gut, to produce therapeutic proteins or to display antigens on their surface to stimulate immune response. These studies have engineered bacteria as treatments for inflammation, diabetes, cancer and infection and have been the subject of several recent literature reviews (41–44). Ideally, an engineered bacterium designed to detect and treat disease should be capable of sensing one or more disease-related or environmental signals and processing the information, before producing a therapeutic compound in response (Fig. 1.1). The incorporation of such disease-responsive capabilities into engineered bacteria enables their use as diagnostic recording tools. For therapeutic applications, making drug production dependent on the detection of a disease signal may lessen any side-effects of the therapy on the host as well as the metabolic burden on the engineered strain. Despite the prospective advantages presented by disease-sensing capabilities, the majority of preclinical and clinical studies with engineered gut bacteria have focused primarily on the expression of therapeutic proteins, either constitutively or by induction with exogenously-supplied compounds. There are several reasons for this, including that 1) even constitutive production of a biologic compound in the gut may represent a significant advantage over oral, intravenous or intramuscular administration, 2) constitutive expression is simpler to engineer than a coordinated response to physiological or pathological signals, and 3) there are relatively few characterized bacterial systems that are known to respond specifically to human disease states.
This dissertation seeks to address the latter two points, focusing on signal sensing and processing capabilities in engineered bacteria. Specifically, this work describes the engineering of commensal bacteria as biosensors of the in vivo gut environment, as well as new mechanisms for constructing genetic circuits that can accurately detect and record multiple recurrences of a signal.

1.3 Bacterial sensors for the intestinal environment

As mentioned above, the majority of bacteria engineered to interface with mammalian systems have been designed to produce a protein either constitutively or in response to an administered inducer compound. However, a few types of synthetic circuits have used bacterial one- and two-component systems to sense signals that may be relevant to mammalian physiology and pathology. One-component systems typically include transcription factors that are directly activated by a specific compound in the cytoplasm. Two-component systems sense stimuli using a kinase with a periplasmic or extracellular ligand-binding domain, and this kinase then transmits the signal by phosphorylating a cytosolic transcription factor (Fig. 1.2). Sensing circuits can be largely classified into biological sensors, which sense small molecules or peptides linked to infection or inflammation (summarized in Table 1.1), and environmental sensors, which sense environmental conditions such as oxygen, temperature and pH.
Figure 1.2: Varieties of sensors used in bacterial synthetic circuits designed to interface with mammalian systems. One-component and two-component systems have been employed in bacteria to construct circuits designed to sense pathogens and inflammation (see Table 1.1), as well as environmental signals inside the host. For simplicity, a single membrane is shown, and any transporters/permeases have been omitted. Triangles represent inducer molecules. R = one-component receptor protein; HK = histidine kinase; RR = response regulator.
TABLE 1.1 Studies with bacteria engineered to detect compounds associated with infection or inflammation.

<table>
<thead>
<tr>
<th>Species</th>
<th>Input</th>
<th>1- or 2-Component</th>
<th>Responding Factor</th>
<th>In Vivo Experiment</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td>P. aeruginosa 3OC12HSL</td>
<td>one</td>
<td>LasR</td>
<td>No</td>
<td>(45), (46), (47)</td>
</tr>
<tr>
<td>E. coli</td>
<td>P. aeruginosa 3OC12HSL</td>
<td>one</td>
<td>LasR</td>
<td>Yes</td>
<td>(48)</td>
</tr>
<tr>
<td>E. coli</td>
<td>V. cholerae CAI-1</td>
<td>two</td>
<td>CqsS/LuxU</td>
<td>No</td>
<td>(49), (50)</td>
</tr>
<tr>
<td>L. lactis</td>
<td>E. faecalis cCF10</td>
<td>one</td>
<td>PrgX</td>
<td>No</td>
<td>(51)</td>
</tr>
<tr>
<td>L. reuteri</td>
<td>S. aureus AIP-1</td>
<td>two</td>
<td>AgrC/AgrA</td>
<td>No</td>
<td>(52)</td>
</tr>
<tr>
<td>E. coli</td>
<td>C. albicans HPA</td>
<td>one</td>
<td>HpaA</td>
<td>No</td>
<td>(53)</td>
</tr>
<tr>
<td>L. lactis</td>
<td>V. cholerae CAI-1</td>
<td>two</td>
<td>hybrid CqsS-NisK/NisR</td>
<td>Yes</td>
<td>(54)</td>
</tr>
<tr>
<td>E. coli</td>
<td>fucose</td>
<td>one</td>
<td>FucR</td>
<td>Yes</td>
<td>(55)</td>
</tr>
<tr>
<td>E. coli</td>
<td>tetrathionate</td>
<td>two</td>
<td>TrrR/TrrS</td>
<td>Yes</td>
<td>(56)</td>
</tr>
<tr>
<td>E. coli</td>
<td>tetrathionate</td>
<td>two</td>
<td>Sbal195_3859/8</td>
<td>Yes</td>
<td>(57)</td>
</tr>
<tr>
<td>E. coli</td>
<td>thiosulfate</td>
<td>two</td>
<td>Shal_3128/9</td>
<td>Yes</td>
<td>(57)</td>
</tr>
<tr>
<td>E. coli</td>
<td>nitric oxide</td>
<td>one</td>
<td>NorR</td>
<td>No</td>
<td>(58)</td>
</tr>
<tr>
<td>E. coli</td>
<td>nitrate &amp; nitrite</td>
<td>two</td>
<td>NarX/NarL and NarQ/NarP</td>
<td>No</td>
<td>(59)</td>
</tr>
<tr>
<td>E. coli</td>
<td>heme</td>
<td>one</td>
<td>HrtR</td>
<td>Yes</td>
<td>(60)</td>
</tr>
</tbody>
</table>

1.3.1 Pathogen-sensing circuits

Several varieties of sensing circuits have been created to detect and respond to compounds produced by pathogenic bacteria in the intestine. Many of these systems have harnessed the pathogens’ native quorum-sensing pathways to enable an engineered commensal strain to respond to the pathogen’s presence.

*Pseudomonas aeruginosa* is a multi-drug-resistant opportunistic pathogen, which can spread from the intestine into the bloodstream and lungs in immunocompromised patients (61, 62). Several circuits have been constructed to detect the presence of *P. aeruginosa* based on the binding of its quorum-sensing compound, 3OC12HSL, to a LasR activator in *E. coli*. In one circuit, *E. coli* was engineered to express the *Pseudomonas*-specific bactericidal pyocin S5, along with the *E. coli*-specific E7 lysis protein (45) in response. When co-cultured with *P. aeruginosa*, this *E. coli* produced pyocin and released it via lysis, inhibiting *P. aeruginosa* growth in liquid
culture, as well as biofilm formation (45). A later study used a different pyocin with an added secretion tag, eliminating the need for lysis for delivery (46). Further improvements to the system included the addition of 3OC_{12}HSL-dependent chemotaxis, enabling the engineered strain to migrate toward the pathogen, combined with the deployment of anti-biofilm and antimicrobial peptides (47). Most recently, a 3OC_{12}HSL-induced circuit producing bacteriocin, an anti-biofilm enzyme and E. coli lysis protein, was integrated into the probiotic strain E. coli Nissle 1917 (48). This engineered strain was tested in a nematode model of P. aeruginosa infection and shown to increase nematode survival. When mice infected with P. aeruginosa were treated with the engineered and wild-type Nissle 1917 strains, the engineered strain reduced the pathogen load more effectively than its wild-type counterpart, though no survival effect was measured. This effect was further magnified when mice were treated with the engineered strain prophylactically prior to P. aeruginosa infection (48).

Another E. coli strain was engineered to respond to Vibrio cholerae by sensing V. cholerae autoinducer 1 (CAI-1) via the V. cholerae two-component sensor CqsS/LuxU (49). This engineered strain produced a reporter in the presence of V. cholerae supernatant in vitro, although its dynamic range was fairly low. This circuit was further expanded to constitutively produce an artilysin with activity against V. cholerae, as well as a lysis protein induced by CAI-1. Supernatant from E. coli that had been grown in V. cholerae supernatant showed a killing effect when subsequently applied to V. cholerae, though no co-culture experiments were conducted (50).

In addition to E. coli, lactic acid bacteria have also been used as a platform for pathogen-sensing circuits. Lactococcus lactis was engineered to detect the Enterococcus faecalis peptide cCF10 and to produce bacteriocins in response, controlled by the E. faecalis PrgX repressor (51).
The engineered *L. lactis* showed a modest inhibitory effect in co-culture with cCF10-producing *E. faecalis* vs. co-culture with a non-cCF10-producing *E. faecalis* strain. However, induction with additional cCF10 resulted in a $10^4$-fold improvement in *E. faecalis* inhibition, suggesting that improving the circuit’s sensitivity to cCF10 would likely increase its potency (51). Another *in vitro* study used *Lactobacillus reuteri* to detect autoinducer peptide-I (AIP-I) of *Staphylococcus aureus* via the *S. aureus* two-component system AgrC/AgrA (52). In the engineered *L. reuteri*, AgrC and AgrA were constitutively expressed, while a reporter gene was placed under the control of the P3 promoter, which is typically activated by phosphorylated ArgA. Although the strain displayed the inverse behavior of what was expected (possibly due to interference from a native transcription factor), the reporter circuit was able to detect AIP-I in supernatant from *S. aureus* cultures (52).

Bacterial circuits have also been built to respond to non-bacterial pathogens, as demonstrated with *E. coli* strain engineered to sense hydroxyphenylacetic acid (HPA) produced by *Candida albicans* (53). In immunocompromised patients, *C. albicans* can penetrate the intestinal epithelium, a process which requires it to assume filamentous form (63). A native *E. coli* HPA transporter (HpaX) and HPA-activated transcription factor (HpaA) were used to control expression of a filamentation-inhibiting compound, cis-2-dodecenoic acid. This engineered *E. coli* strain was able to inhibit filamentation in *C. albicans* in co-culture (53).

Besides harnessing existing one- and two-component systems for pathogen sensing, novel hybrid sensors have also been constructed. Recently, an *in vivo* study engineered *L. lactis* with a hybrid histidine kinase to detect *V. cholerae* in the intestine by sensing *V. cholerae* CAI-1 (54). This hybrid receptor consisted of the ligand-binding domain of the *V. cholerae* CAI-1 receptor, CqsS, and the kinase domain of the *L. lactis* nisin receptor, NisK from the NisK/NisR two-
component system. The hybrid sensor allowed CAI-1 to activate the response regulator NisR, inducing a reporter gene. Administration of the sensor strain and *V. cholerae* to mice demonstrated that *V. cholerae* could be detected using a stool-based readout of this reporter. This study also examined the natural inhibitory effects of *L. lactis* on *V. cholerae* colonization. Curiously, the natural protective effect of wild-type *L. lactis* against *V. cholerae* was reduced in the engineered *L. lactis* strain—presumably due to the metabolic burden of the engineered circuit resulting in reduced acidification—illustrating the importance of burden and context considerations for engineered strains (54).

In addition to directly sensing pathogens, bacteria can also be used to sense changes in their local environment resulting from systemic infection. The activation of toll-like receptors by pathogen-associated molecular patterns (PAMPs) in systemic circulation leads to fucosylation of the small intestine epithelium, which can increase tolerance to intestinal pathogens (55). This fucose is then cleaved by intestinal bacteria and the free fucose can be detected by *E. coli*. One study used a fluorescent reporter under the control of the native *E. coli* fucPIK UR promoter to sense the presence of fucose in the intestine after systemic administration of various PAMPs (55).

### 1.3.2 Inflammation-sensing circuits

Bacterial circuits have also been designed to sense molecules associated with an inflammatory response. A commensal murine *E. coli* strain was engineered to detect tetrathionate (56), which is produced during inflammation as a result of thiosulfate oxidation by reactive oxygen species released from the gut epithelium (64). The two-component system TtrR/TtrS from *Salmonella enterica* Typhimurium (S. Typhimurium), which utilizes tetrathionate for anaerobic respiration (64), was adapted for this purpose. Upon sensing tetrathionate, this engineered *E. coli* strain activates a memory switch driven by the *ttrBCA*
promoter. This strain was able to detect and record intestinal inflammation in several mouse models, including IL-10 knockout mice and mice infected with *Salmonella*. Furthermore, this strain was able to stably colonize mice and maintain a functioning memory circuit for at least six months, demonstrating the low burden of this synthetic circuit (56). In another study, *E. coli* Nissle 1917 strains were engineered to sense tetrathionate, as well as thiosulfate, another putative biomarker of inflammation (57). These strains were designed to produce a fluorescent reporter in response to activation of two-component systems sourced from *Shewanella baltica* and *Shewanella halifaxensis*. Both systems responded specifically to their corresponding inducer *in vitro*. In a dextran sulfate sodium-treated mouse model of inflammation, the thiosulfate sensor produced a response, while the tetrathionate sensor did not, though this may be because this tetrathionate sensor was less sensitive than the TtrR/TtrS system described above (56, 57). Further examination is needed to determine the feasibility of using tetrathionate and thiosulfate as inflammatory biomarkers in the clinic.

Host inflammation is also associated with nitrogen oxides such as nitric oxide, nitrate and nitrite (65). In one study, *E. coli* were engineered to sense nitric oxide via the NorR-P$_{norV}$ system and to trigger a memory switch as a result (58). The sensor bacteria responded in co-culture with an explant model of mouse ileal cells, and this response was greatly reduced when the ileal cells were treated with an inhibitor of nitric oxide synthase—demonstrating an *in vitro* proof-of-concept for nitric oxide sensing (58). Another study used the promoter of the *E. coli yeaR-yoaG* operon linked to a memory switch to construct sensors that detect nitrate and nitrite and showed their ability to respond in *in vitro* culture (59).

While not directly sensing inflammation, another circuit in *E. coli* was recently engineered to detect intestinal bleeding via the presence of heme (60). A promoter under the
control of the *L. lactis* HrtR repressor was engineered to control the expression of a luciferase reporter. The heme-sensing strain was shown to respond in mice when intestinal bleeding was induced by the administration of indomethacin. In a follow-up experiment, this bacterial sensor was encased in an electronic capsule, which could detect luciferase and wirelessly transmit a signal from inside the host. This capsule was deposited into the stomach of pigs and responded when a blood and bicarbonate solution was administered to the pigs orally (60).

1.3.3 Sensing environmental factors

Along with detecting pathogens and inflammation, sensors that respond to environmental conditions encountered in the gut and elsewhere in the body may be useful for constructing therapeutic circuits.

The ability to sense oxygen concentration is of great interest for programming therapeutic bacteria for deployment in the intestine, as well as in solid tumors. The gastrointestinal tract exhibits a decreasing gradient of oxygen concentration both longitudinally and radially, with the partial pressure of oxygen in the lumen typically below 10 mmHg (66). In the gut, oxygen concentration might be used to allow production of a therapeutic compound at an intended location or to trigger a kill switch that would result in death of the engineered bacteria outside their designated environment. Solid tumors also typically have lower oxygen concentrations than healthy tissue, with some tumor microenvironments registering less than 2.5 mmHg (67). Because of this, natural and engineered facultative anaerobic bacteria capable of colonizing tumors have been explored as potential cancer therapies (67). *E. coli* has a broad range of genes that are differentially expressed in anaerobic vs. aerobic conditions, many of which are under the direct or indirect control of the global regulator FNR (68). To target tumors, an anaerobically-inducible circuit in *E. coli* was designed to produce an invasin gene from *Yersinia*
*pseudotuberculosis*, which allows the strain to adhere to and invade mammalian cells (69).

Production of the invasin was controlled by an anaerobically-inducible promoter from the *E. coli* *fdhF* gene. The ability of this strain to invade mammalian cells in anaerobic conditions was confirmed by an *in vitro* invasion assay (69). Another study used the promoter from the *E. coli* *nirBDC* operon to construct an anaerobically-induced circuit in *S. Typhimurium* (70). This circuit produced an apoptosis-inducing ligand in low-oxygen conditions, and systemic administration of this strain effectively reduced the size of tumors in a mouse model of melanoma. Furthermore, the reduction in tumor volume was more pronounced with the engineered ligand-producing strain than with administration of a control strain combined with daily injections of the ligand (70).

Temperature- and pH-sensing systems are also of interest, particularly those that can respond differentially to host vs. ambient temperature and to the pH gradient encountered in the gastrointestinal tract. An AND logic gate was constructed in *E. coli*, which required sensing a combination of cold temperature (via the *cspA* promoter) and low pH (via the *asr* promoter) to produce a fluorescent output (71). P_{asr} and P_{cspa} were used to produce the transcription factor, InvF, and chaperone, SicA from *S. Typhimurium*. Together, these allowed expression of a reporter under the control of P_{SicA}. An inverter was further added to this circuit to create the corresponding NAND gate (71). Other examples of temperature sensors include two studies that used temperature-sensitive circuits to control kill switches for containment of bacteria inside the host. One study engineered *E. coli* to constitutively produce the toxin CcdB but used a temperature-dependent promoter—controlled by a mutant of the *S. Typhimurium* autorepressor, TlpA—to allow expression of an antitoxin at temperatures above 36 °C (72). This engineered strain was administered to mice by gavage and recovered from fecal samples. Growth
of the recovered bacteria at 25 °C resulted in a 10³-fold decrease in survival compared to continued growth at 37 °C (72). Another kill switch was designed to constitutively express a low level of antitoxin but used the cspA promoter to trigger production of excess CcdB at low temperature (73). With this improved kill switch, growth of bacteria recovered from mice at 22 °C resulted in a 10⁵-fold decrease in survival compared to continued growth at 37 °C. This circuit also demonstrated stability after prolonged growth in vitro and remained equally functional after 10 days in the mouse gut. This stability was possibly a result of the low-level antitoxin expression, which may discourage mutation of the toxin gene even with leaky expression of the toxin (73).

1.3.4 Engineering and screening for new biosensors

The construction of host-responsive systems in bacteria is currently limited by the very small set of characterized systems known to respond to specific biomarkers or environmental conditions. Currently, there are just a handful of characterized sensors that have been used to engineer bacteria to interface with mammalian systems, many of which have only been shown to work in vitro. These are limited to sensors responding to 1) small molecules and peptides produced by several pathogens 2) nitrogen- and sulfur-oxides related to inflammation, and 3) environmental factors such as oxygen, temperature and pH. Thus, there is a need for expanding the arsenal of biosensors that can be employed in the design of engineered bacterial diagnostics and therapeutics. One approach involves the design of new sensors based on fusion proteins. For instance, an existing protein domain known to bind to a biomarker of interest can be employed in a new context by combining it with another domain that activates the desired response; this can be achieved through engineering one-component (74) or two-component (54) systems. However, this technique is also limited by our existing knowledge of sensor–biomarker interactions.
An alternative approach involves mining of bacterial genomes to discover new sensors that respond specifically in conditions of interest. In particular, native microbiota, which are adapted to the host environment, may prove to be a rich source of systems that can be employed as biosensors. In addition to the sensors described so far, there is evidence that bacteria are capable of responding to a variety of host factors including: neurotransmitters and hormones such as serotonin (75), γ-aminobutyric acid (76), epinephrine and norepinephrine (77); endogenous opioids such as dynorphin (78); and cytokines including IL-1β (79), TNF-α (80) and IFN-γ (81).

Discovering and harnessing such natural sensors requires characterizing their activity. Most systems-level studies of bacterial responses in the host have used techniques such as transcriptomics (82) and proteomics (83, 84), which can provide an instantaneous snapshot of bacterial behavior. One disadvantage is that to monitor bacterial behavior while inside the host, these techniques require invasive sampling. Furthermore, in the absence of a specific filtering or selection method, these approaches may not detect low-abundance signals. Finally, any responsive sensor elements they identify from their natural context would likely require further tuning to function in the context of a specific synthetic circuit.

One method that stands in contrast to these techniques is recombinase-based in vivo expression technology (RIVET) (85), which can be used to screen for bacterial promoters that are active inside the host by using a recombinase-based memory switch to record transient promoter activity. RIVET screening is based on the construction of libraries comprising genomic DNA fragments inserted upstream of a promoter-less recombinase gene. If expressed during transit through the gut, the recombinase removes a selective marker at another specific locus; thus, any strains that have lost their selective marker after gut transit are likely to contain DNA fragments
that are transcriptionally active in the gut environment. While RIVET has been used to study microbial gene expression in a wide range of contexts (86, 87), most RIVET studies in mammals have examined bacterial pathogenicity mechanisms—most notably in *V. cholerae*—by screening for (85, 88) or tracking temporal expression of (89) virulence genes. While RIVET initially had false positive rates of over 60% (85), some improvements were made to increase the method’s efficiency (89–91). These screening methods demonstrate the promise of searching bacterial genomes for sensors that can respond to the host environment. The second chapter of this dissertation presents a new, highly accurate, systems-level approach for discovering and fine-tuning bacterial biosensors for the construction of *in vivo*-responsive circuits.

### 1.4 Circuits for processing & recording signals

In addition to producing a response based on the detection of a single signal, more complex engineered circuits can be programmed to respond based on a variety of inputs. To the best of our knowledge, the only current example of a bacterial circuit that combines multiple physiologically-relevant inputs is the pH–temperature AND gate described above (71). However, the field of synthetic biology has made great strides in constructing general-use circuits that can process and record information in ever more complex ways, which may see application in future engineered diagnostics and therapeutics. These can be generally classified into circuits that carry out logic operations (producing outputs based on specific combinations of inputs) and information storage (storing a record of inputs over time)—although many circuits possess aspects of both categories.
The first logic circuits took advantage of basic transcriptional regulation, such as OR and NOR gates based on combinations of promoters (92), as well as AND gates based on two-part transcription factor–chaperone activators (93). More recent logic circuits have been built using recombinases, which provide a digital output along with the ability to make a record of the signal (94). However, unlike transcription factor-based logic, activation of recombinase-based circuits is often unidirectional. While reversible recombinases may be implemented (95), the practicality of such an approach depends on the application. More recently, logic circuits have also been implemented using programmable RNA-based ribocomputing (96).

Biological recording circuits can be classified into analog and digital varieties. Analog recorders are typically used to track the intensity of a stimulus applied over time. Recorders such as SCRIBE (97) and CAMERA (98) accumulate mutations or deletions in the bacterial genome or on plasmids which can be read out through population-level sequencing or by the resulting alteration of a reporter gene’s function. Another analog system used variations in single-cell recombinase states across a population of cells to produce a population-level, analog record of the order and duration of two applied stimuli (99). Yet another analog recording strategy used the acquisition of short CRISPR spacer sequences via the Cas1–Cas2 complex—either through exogenous introduction of spacer oligos at defined intervals (100) or through stimulus-driven manipulation of plasmid copy numbers (101), resulting in incorporation of different ratios of spacer sequences into a CRISPR array.

Rather than continuously logging signal intensity, digital recorders are used to record when a signal crosses a certain threshold. The first digital recording circuit—the toggle switch (102)—is considered to be one of the studies that launched the field of synthetic biology. This switch used two repressors that repress each other’s production, which resulted in two stable
states. Switching from one state to the other could be achieved through inhibition of the active repressor with the corresponding inducer molecule. Like the toggle switch, the bacteriophage \( \lambda \)-based memory element (103) contained a bistable switch based on two mutually-repressing proteins. From a memory-off state, corresponding to constant production of \( \lambda \) CI, the switch could be flipped on via a separate trigger element, which produced exogenous \( \lambda \) Cro, corresponding to the memory-on state. Bacteria containing this circuit were shown to colonize the mouse intestine and could record the presence of an inducer that was administered via the drinking water (103). This same circuit was later adapted to sense tetrathionate in the gut, and these memory bacteria were further shown to maintain functional and genetic stability in mice for over six months without requiring antibiotic maintenance (56), demonstrating this circuit’s suitability for long-term \textit{in vivo} monitoring. Digital recorders have also been built using recombinases, which can make permanent or reversible changes in DNA sequences to store information. Arrays of recombinase sites can be used to encode a single signal as a combination of multiple bits (104) or to allow recording the order in which several different signals are applied (105).

A specific type of digital recorder is the pulse counter, which records the number of discrete applications of a single type of stimulus. To date, two pulse-counting circuits have been constructed (106). One circuit is controlled by riboswitches, which are each successively activated with pulses of arabinose inducer. Each arabinose pulse opens the next riboswitch and allows the translation of the unique RNA polymerase required for transcription of the next mRNA/riboswitch in the sequence. The second circuit is based on linear arrangement of DNA elements, each containing a recombinase gene and terminator, and each flanked by recombinase sites. Each pulse of arabinose enables the inversion of the next DNA element in the sequence,
which disables its terminator and allows transcription of the next recombinase next arabinose pulse (106). While these circuits produced varying responses with subsequent activations of a single promoter, both could only count arabinose pulses that were applied for a specific time and at a specific concentration. Outside a narrow window of pulse duration and concentration, a single strong stimulus application might make the counters record multiple counts, making them impractical for real-world deployment.

The third chapter of this dissertation describes the construction of pulse counters with counting mechanisms incorporating negative feedback. These circuits introduce several improvements to previously-constructed pulse counters, including the ability to record only a single count for a single stimulus pulse, even with pulses applied for over 24 hours.

### 1.5 Aims of this study

The goal of this dissertation is to expand the capabilities of synthetic biology for the purpose of constructing synthetic bacterial circuits to interface with mammalian hosts. This dissertation focuses on two underexplored areas of need: Chapter 2 describes new methods for the discovery of new sensors responding to the *in vivo* host environment. Chapter 3 describes the construction of functional pulse counting circuits for recording discrete stimulus pulses of virtually unlimited magnitude and concentration. The concluding chapter discusses considerations for real-world deployment of engineered bacteria from both an engineering and a societal angle.
Chapter 2: A synthetic memory circuit enables systems-level biosensor discovery at the host–microbe interface
Preface

The project described in this chapter owes a lot to both fundamental work and continuing support from many labmates. David Riglar has been a wonderful mentor throughout the duration of this project. He and Johannes Bues initiated the project, constructed the MG1655 library and conducted the first one-day mouse screen. Shannon Nangle & Disha Trivedi constructed the Nissle 1917 library. Nabeel Nissar, Disha Trivedi & Matt Niederhuber assisted with in vitro characterization. Mara Inniss constructed λ cIDN, which was widely used in this project. Adam Riesselman and Michiel Karrenbelt wrote promoter delineator and protein homology comparison scripts which were used in the design and construction of the NGF-1 library. Bryan Hsu, Andrew Verdegaal, and Joseph Paulson provided helpful input. It has been a pleasure to collaborate with such a talented group of people, and I am grateful for your contributions and perspectives. The majority of this chapter is adapted from a submitted manuscript, posted as a preprint on bioRxiv:


2.1 Abstract

The composition and function of the gut microbiota are strongly associated with human health, and dysbiosis is linked to an array of diseases, ranging from obesity and diabetes to infection and inflammation. Engineering synthetic circuits into gut bacteria to sense, record and respond to in vivo signals is a promising new approach for the diagnosis, treatment and prevention of disease. Here, we repurpose a synthetic bacterial memory circuit to rapidly screen
for and discover new in vivo-responsive biosensors in commensal gut Escherichia coli. We develop a pipeline for rapid systems-level library construction and screening, using next-generation sequencing and computational analysis, which demonstrates remarkably reliable identification of responsive biosensors from pooled libraries. By testing both genome-wide and curated libraries of potential biosensor triggers—each consisting of a promoter and ribosome binding site (RBS)—and using RBS variation to augment the range of trigger sensitivity, we identify and validate triggers that selectively activate our synthetic memory circuit during transit through the gut. We further identify biosensors with increased response in the inflamed gut through comparative screening of our libraries in healthy mice and those with intestinal inflammation. Our results demonstrate the power of systems-level screening for the identification of novel biosensors in the gut and provide a platform for disease-specific screening using synthetic circuits, capable of contributing to both the understanding and clinical management of intestinal illness.

2.2 Introduction

Recent advances in our understanding of both the human microbiota and biological engineering techniques have created myriad possibilities for the development of synthetic microbes for in vivo clinical applications (44, 107). Bacteria residing in the gut are uniquely positioned to monitor a variety of host, microbial, and environmental factors and to respond to changes in intestinal homeostasis. Engineered gut bacteria also offer the potential for in vivo production and delivery of therapeutics (44).
Environment- and disease-responsive functions, which could minimize both the metabolic burden of engineered systems on the bacteria and off-target effects on the patient, offer exciting prospects for clinical applications. To this end, recent in vivo approaches have developed sensors responding to inflammation (56, 57), intestinal bleeding (60), and pathogen quorum-sensing systems (48, 54). However, the construction of disease-responsive circuits in bacteria has been hindered by the limited number of characterized bacterial systems that can be reliably employed as sensors.

Mining the genomes of native gut bacteria is a promising approach for discovering new sensors that respond under conditions of interest, such as in the healthy or diseased gut. To date, these efforts have largely relied on transcriptome sequencing and proteomics of fecal samples. However, to obtain an instantaneous snapshot of bacterial behavior inside the gut using these techniques, invasive sampling is required (i.e., colonoscopy and biopsy). Furthermore, transient or low-abundance signals may not be detected, and any responsive genetic elements identified with these techniques may not function predictably when employed in synthetic circuits. Approaches such as in vivo expression technology (IVET) and recombinase-based IVET (RIVET) have also been used to track in vivo-expressed genes non-invasively, but detect only constitutive expression (for IVET) and may have high false-positive rates (86). Nevertheless, these technologies show the potential for systems-level approaches to interrogate the behavior of the microbiota.

We have previously developed an approach for non-invasive measurement of bacterial responses in the gut, based on a robust synthetic memory circuit, which records environmental stimuli via a transcriptional trigger (56, 103). When activated, the trigger turns on a memory switch, which can retain the memory-on state for over a week in the gut (103). After the bacteria
pass through the host, their memory state can be determined via reporter gene expression, enabling non-invasive readout of transient signals within the gut. The circuit can maintain functional and genetic stability during six months’ colonization of the mouse gut, demonstrating its suitability for longitudinal studies and its potential to support the development of stable, engineered biosensors for \textit{in vivo} deployment (56).

Here, we adapt this memory circuit for parallel, high-throughput screening of hundreds of potential triggers. We apply this method to identify new biosensors responding specifically to the \textit{in vivo} gut environment. Through comparison between healthy mice and those suffering from inflammation, we also identify triggers that respond differentially during disease. Together, these results provide a platform for \textit{in vivo} non-invasive biosensor discovery and longitudinal testing.

\section*{2.3 Results}

\subsection*{2.3.1 Bacterial memory as a biosensor screening tool}

To enable screening of new potential biosensors in parallel, we modified our previously-developed \textit{E. coli} memory circuit, which is based on the phage \textit{\lambda} lysis–lysogeny switch (Fig. A.1A) (103). This modified circuit is referred to as the high-throughput memory system (HTMS) (Fig. 2.1A). Both the original memory circuit and the HTMS consist of a trigger—based on a transcriptional promoter activated in the presence of a certain stimulus—and a memory switch. The memory-on and memory-off states of the switch correspond to the mutually-repressive proteins Cro and CI, respectively. In the memory-off state, CI is produced under the control of $P_{RM}$, and represses $cro$, controlled by $P_{R}$. In the original circuit, when the trigger promoter is
activated by a stimulus, Cro is produced, repressing $P_{RM}$ and $cI$. After dilution of the remaining CI protein through cell division, $P_R$ becomes active, and the circuit transitions to the memory-on state, which also produces a $\beta$-galactosidase (LacZ) reporter.

Figure 2.1: Design of a High-Throughput Memory System (HTMS). (A) HTMS circuit design in memory-off and memory-on states. (B) Comparison of memory element induction with $cro$ and $cIDN$ triggers illustrates differences in induction dynamics. Control and memory strains with $P_{tet}$ triggers (PAS132, PAS133, PAS807, PAS808) were grown in liquid media, then spotted on indicator plates, each with or without aTc induction (100 ng/ml). Images courtesy of Matthew Niederhuber. (C) Selection of memory-on HTMS with spectinomycin. Memory-off, memory-on, and 50-50 mixed cultures of PAS810 were plated on indicator plates with or without spectinomycin (inset photo). Graph shows comparison of CFU counts between +spectinomycin and -spectinomycin plates. Error bars represent SE of eight biological replicates (for 0% and 100%) and five biological replicates (for 50%).
One key modification for the HTMS is the triggering of memory using a dominant-negative mutant of the cl gene (cIDN), instead of cro used in the original trigger. When the trigger promoter is activated by a stimulus, CIDN monomers, which have an N55K mutation in the DNA-binding region (108) dimerize with wild-type (WT) CI monomers expressed in the memory-off state, creating heterodimers that are deficient in DNA binding. This leads to derepression of Pr and transition to the memory-on state. As with the CI used in the memory switch, CIDN carries a mutation to prevent RecA-mediated cleavage (ind-) (109).

Use of CIDN in the trigger ensures that there is no delay of switching to the memory-on state in the case of high, or constant, expression of the trigger promoter. To test this, a Ptet trigger driving cIDN or cro was integrated into E. coli K-12 MG1655 and NGF-1 strains containing a memory switch. When grown in the presence of a high concentration (100 ng/ml) of anhydrotetracycline (aTc) inducer, cIDN-triggered strains showed switching to the memory-on state, while cro-triggered strains switched only after a subsequent period of growth in the absence of aTc (Fig. 2.1B; see also section 3.3.1).

The original memory circuit expresses a lacZ reporter gene for screening on indicator plates (103). To analyze pooled libraries containing many strains with varied trigger promoters, the HTMS also expresses a spectinomycin-selectable resistance gene (aadA) in the memory-on state.

This antibiotic-selectable memory maintains response characteristics similar to the original memory switch. To test this, a Ptet trigger driving cIDN was integrated into strains containing lacZ (original) or aadA+lacZ (HTMS) memory switches, creating PAS809 and PAS810, respectively (see Table 2.1 for list of strains used in this chapter). Strains were induced by aTc (0-100 ng/ml) and the response quantified by plating cultures on indicator plates.
containing 5-bromo-4-chloro-3-indolyl-β-d-galactopyranoside (X-gal), which turns blue in the presence of LacZ, indicating a memory-on state (Fig. A.1B). Both strains responded similarly to aTc (original memory EC50: 4.1-4.6 ng/ml, 95% CI; HTMS EC50: 4.0-4.3 ng/ml, 95% CI), confirming the circuit’s modularity to additional reporters in the memory-on state.

**TABLE 2.1** Key memory strains used in this study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>E. coli strain</th>
<th>Memory</th>
<th>Trigger Promoter/RBS</th>
<th>Trigger Gene</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAS132</td>
<td>K-12 MG1655</td>
<td>original</td>
<td>tet</td>
<td>cro</td>
<td>(103)</td>
</tr>
<tr>
<td>PAS133</td>
<td>NGF-1</td>
<td>original</td>
<td>tet</td>
<td>cro</td>
<td>(103)</td>
</tr>
<tr>
<td>PAS807</td>
<td>K-12 MG1655</td>
<td>original</td>
<td>tet</td>
<td>cIDN</td>
<td>this study</td>
</tr>
<tr>
<td>PAS808</td>
<td>NGF-1</td>
<td>original</td>
<td>tet</td>
<td>cIDN</td>
<td>this study</td>
</tr>
<tr>
<td>PAS809</td>
<td>NGF-1</td>
<td>original</td>
<td>tet</td>
<td>cIDN</td>
<td>this study</td>
</tr>
<tr>
<td>PAS810</td>
<td>NGF-1</td>
<td>HTMS</td>
<td>tet</td>
<td>cIDN</td>
<td>this study</td>
</tr>
<tr>
<td>PAS811</td>
<td>NGF-1</td>
<td>HTMS</td>
<td>-</td>
<td>-</td>
<td>this study</td>
</tr>
<tr>
<td>PAS812</td>
<td>NGF-1</td>
<td>HTMS</td>
<td>fabR</td>
<td>cIDN</td>
<td>this study</td>
</tr>
<tr>
<td>PAS813</td>
<td>NGF-1</td>
<td>HTMS</td>
<td>ydiL</td>
<td>cIDN</td>
<td>this study</td>
</tr>
<tr>
<td>PAS814</td>
<td>NGF-1</td>
<td>HTMS</td>
<td>ydjL</td>
<td>cIDN</td>
<td>this study</td>
</tr>
<tr>
<td>PAS815</td>
<td>NGF-1</td>
<td>HTMS</td>
<td>-</td>
<td>cIDN</td>
<td>this study</td>
</tr>
<tr>
<td>PAS816</td>
<td>NGF-1</td>
<td>HTMS</td>
<td>ynfE15</td>
<td>cIDN</td>
<td>this study</td>
</tr>
<tr>
<td>PAS817</td>
<td>NGF-1</td>
<td>HTMS</td>
<td>yeaRWT</td>
<td>cIDN</td>
<td>this study</td>
</tr>
<tr>
<td>PAS818</td>
<td>NGF-1</td>
<td>HTMS</td>
<td>ynfEWT</td>
<td>cIDN</td>
<td>this study</td>
</tr>
<tr>
<td>PAS819</td>
<td>NGF-1</td>
<td>HTMS</td>
<td>ynfE17</td>
<td>cIDN</td>
<td>this study</td>
</tr>
<tr>
<td>ADN149</td>
<td>NGF-1</td>
<td>HTMS</td>
<td>A9Z04_04088</td>
<td>cIDN</td>
<td>this study</td>
</tr>
</tbody>
</table>

The HTMS allows faithful selection of memory-on colonies with spectinomycin treatment. Plating of fully memory-off, fully memory-on, and 50-50 mixed cultures of PAS810, on indicator plates with and without spectinomycin further demonstrated that all spectinomycin-selected colonies were also lacZ positive (Fig. 2.1C). Spectinomycin did not yield false-positive results by inducing memory switching (fully memory-off: 0% ± 0% SE, n = 8), nor excessive false-negative results through inhibition of memory-on bacterial growth (fully memory-on: 93.0% ± 4.2% SE, n = 8; 50-50 mix: 49.7% ± 1.5% SE, n = 5) (Fig. 2.1C). Together
these results demonstrate the ability of the HTMS to measure biosensor response and allow selection for downstream pooled analyses.

2.3.2 Biosensor library construction

To build biosensor libraries for genomic integration, we adapted a Tn7 transposon genome insertion plasmid (110) for rapid Golden Gate assembly (12) of bacterial promoters upstream of the cIDN trigger and insertion into the genome of memory bacteria (Fig. 2.2A and A.2). The modularity of this cloning strategy allows for adjustment of trigger sensitivity through incorporation of ribosomal binding site (RBS) variants, which vary the translation rate of mRNA transcripts (Fig. 2.2A). To test this concept, triggers consisting of a Ptet promoter combined with nine synthetic RBS sequences—previously demonstrated to vary widely in their translation strength (111) (Fig. 2.2A)—were constructed and inserted into the genome of HTMS bacteria, and the HTMS response to varying concentrations of aTc (0-100 ng/ml) was characterized (Fig. 2.2B). The RBS variants differed in their extent of memory induction at 0.1-10 ng/ml aTc (EC50 ranging from 0.5 to 4.1 ng/ml for responsive strains), illustrating our ability to tune trigger sensitivity.
We explored two approaches for generating biosensor libraries: 1) genome-wide collections of trigger promoters that would enable screening of a bacterium’s entire range of transcriptional responses (MG1655 and NGF-1 libraries), and 2) a curated collection of promoters with sensitivity variants aimed at detecting inflammation (Nissle 1917 library). All libraries were assembled into an HTMS-containing E. coli NGF-1 strain (PAS811), as NGF-1 has proven to be an efficient and persistent colonizer in the mouse gut (56, 103, 112).

The MG1655 genome-wide library was sourced from a previously published collection of 1600 unique promoters from E. coli K-12 MG1655 (113). Promoters and their wild-type RBSs were amplified by PCR from this collection, assembled into our transposon plasmid, and integrated as triggers into the genome of PAS811. Because our method focuses on detecting off-to-on sensor transitions, the resultant library was further subsampled by pooling 500 colonies.
that were LacZ-negative under routine *in vitro* culture. Sequencing confirmed the presence of 155 unique strains in this final MG1655 library.

The NGF-1 genome-wide library was built using promoters of genes found specifically in the native gut strain *E. coli* NGF-1, which are not shared with K-12 MG1655 (where less than 90% of the protein sequence of the NGF-1 gene has at least 70% identity to a K-12 MG1655 gene). These genes were further filtered based on NGF-1 genome annotations (112) corresponding to regulatory functions, enzymatic activity and stress response. From these, 384 predicted promoters with their wild-type RBSs were amplified from the NGF-1 genome by PCR and incorporated as triggers in the NGF-1 library. Sequencing confirmed that the assembled NGF-1 library contained 381 unique strains out of 384 total designed constructs.

Our curated Nissle 1917 library was constructed with a subset of promoters sourced from the human probiotic *E. coli* Nissle 1917, which are involved in anaerobic respiration of sulfur- or nitrogen-oxides or nitrate, produced by the gut epithelium during inflammation (65, 114). For each promoter, a trigger with its wild-type RBS, as well as with five different synthetic RBSs (MCD5, MCD10, MCD15, MCD17 and MCD23) (111) was included to tune sensitivity. Throughout this dissertation, strains from this library are referred to by an abbreviation consisting of the first gene of the operon from which their trigger is derived and the number of the synthetic RBS used. For instance, “ynfE15” denotes the trigger consisting of the ynfEFGH promoter with MCD15. Sequencing confirmed that the assembled Nissle 1917 library contained 61 unique strains out of 66 total designed constructs.

### 2.3.3 Parallel analysis faithfully reports biosensor response

To screen for biosensor response, HTMS libraries are exposed to a condition of interest (Fig. 2.3A), and put through a processing, sequencing and analysis pipeline (Fig. 2.3B). After
exposure, HTMS bacteria are recovered and cultured. The initial culture is split into two and back-diluted, and one of the two new cultures is subjected to spectinomycin selection. Following selection, the trigger regions of both cultures are sequenced and analyzed to produce an odds ratio for each trigger promoter in the library, corresponding to that trigger’s memory state. To calculate odds ratios, results are normalized to a positive normalization strain (PAS812 for the MG1655 library and Nissle 1917 library; ADN149 for the NGF-1 library) which remains in a constant memory-on state under in vitro culture conditions (Fig. 2.3B).

To test that pooled library analysis is predictive of the on/off state of HTMS bacteria, the Nissle 1917 library was cultured aerobically in liquid media and analyzed to obtain odds ratios as described above. Concurrently, individual strains from this library were grown on indicator plates to assess each strain’s in vitro memory state directly. Both tests showed strong agreement, with strains that were LacZ-positive also displaying higher odds ratios (Fig. 2.3C). Receiver operating characteristic analysis confirmed efficient distinction between memory-on and memory-off states, with an odds ratio of approximately 0.02 delimiting the boundary (Fig. 2.3D). This confirmed our sequencing method as a reliable indicator of biosensor memory state.
Figure 2.3: Biosensor library screening and analysis. (A) Libraries were constructed as plasmids, integrated into the genome at single copy, and screened as a pool for differential response to growth in control and test conditions.
**Figure 2.3 (continued)** environments. (B) Post-exposure library sample processing, selection for memory-on strains, sequencing and analysis. The consistently memory-on positive normalization strain (PAS812 for the MG1655 library and Nissle 1917 library; ADN149 for the NGF-1 library) is spiked in prior to spectinomycin selection and used for calculation of odds ratios. (C) Calculated odds ratio from *in vitro* pooled growth in LB medium vs. memory state assessed by plating individual library strains on LB agar indicator plates. 44 strains subsampled from the Nissle 1917 library were tested. The on-off odds ratio cutoff used in the subsequent *in vivo* screens (0.02) is indicated by the dotted line. (D) ROC curve for varying odds ratio cutoffs as an indicator of memory state.

### 2.3.4 Differential biosensor response in the healthy mouse gut

To screen for biosensor response to growth within the murine gut, the MG1655 library was administered to specific-pathogen free (SPF) mice by oral gavage (~$10^7$ bacteria/mouse), and fecal samples were collected over one ($n=2$) or seven ($n=3$) days. High library diversity was maintained in both experiments with 92% and 82% of strains identified in gavage samples present at experiment endpoint, respectively (Data Set A.1 and A.2). Analysis of HTMS strains recovered from gavage suspension and fecal samples identified 23 unique strains that responded specifically to growth within the gut (gavage: odds ratio < 0.02; fecal samples: $\geq 1$ timepoint odds ratio $\geq 0.02$ and $p < 0.05$) (Fig. 2.4A and 2.4B; Data Set A.1 and A.2). Five strains (containing ydiL, ydjL, gatY, gcvA and ubiG triggers) were detected in the memory-on state in at least 4 of 5 mice. The two most consistent responders (ydjL and ydiL triggers) were selected for follow-up testing.
Figure 2.4: Library screening and individual sensor testing identifies biosensors of the in vivo gut environment. (A) Screen of MG1655 library in BALB/c mice over one day (left; n = 2) and over seven days (right; n = 3). Odds ratio heat maps of the top 10 hits sorted by percentage of positive timepoints (odds ratio ≥ 0.02 and p < 0.05) over the course of the experiment. Blank spaces on heat maps represent insufficient sequencing coverage. See Data Set A.1 and A.2 for full results. (B) Percentage of positive timepoints (odds ratio ≥ 0.02 and p < 0.05) for all positive hits from two MG1655 library screens (top 10 shown in Fig. 2.4A).
Figure 2.4 (continued) (C) Response of individual strains from MG1655 library in healthy mice. HTMS strains containing triggers ydiL (PAS813), ydjL (PAS814) and an empty trigger (PAS815) were administered as monocultures to BALB/c mice (n = 3). Memory response was assessed by plating of HTMS bacteria recovered from fecal samples on indicator plates. Error bars represent SE. (D) Screen of Nissle 1917 library in C57BL/6J mice (n = 4) over five days. Odds ratio heat map of the top 10 hits sorted by percentage of positive timepoints (odds ratio ≥ 0.02 and p < 0.05) over the course of the experiment. Blank spaces on heat map represent insufficient sequencing coverage. See Data Set A.3 for full results. (E) Percentage of positive timepoints (odds ratio ≥ 0.02 and p < 0.05) for all hits from Nissle 1917 library screen in healthy mice (top 10 shown in Fig. 2.4D). (F) Response of individual strains from Nissle 1917 library in healthy mice. HTMS strains containing the ynfE trigger with MCD15 (PAS816), the yeaR trigger with its WT RBS (PAS817) and the ynfE trigger with its WT RBS (PAS818) were administered as monocultures to C57BL/6J mice (n = 3). Memory response was assessed by plating of HTMS bacteria recovered from fecal samples on indicator plates. Error bars represent SE. (G) Response of the ynfE15 trigger strain (PAS816) to in vitro growth in rich media with and without 20 mM nitrate, with and without oxygen, and in the presence of mouse cecum fluid medium. Memory response was assessed by plating on indicator plates after growth in liquid culture. Error bars represent SE. (n = 7 for -O₂, LB, no inducer condition; n = 3 for all other conditions.)

To validate the response of the ydiL and ydjL triggers during gut transit, memory bacteria containing these triggers (ydiL: PAS813; ydjL: PAS814) or a promoter-less cIDN gene (negative control: PAS815) were administered to SPF mice as monocultures. Fecal samples were collected and analyzed over the subsequent five days. Culture on indicator plates demonstrated an absence of memory activation in all three strains prior to gavage. However, when recovered from fecal samples PAS813 and PAS814 colonies were consistently memory-on, confirming activation.
during gut transit (at Day 2, PAS813: 51% ± 8% SE; PAS814: 30% ± 4% SE; negative control: 0% ± 0% SE; n = 3 per strain) (Fig. 2.4C).

The Nissle 1917 library was also screened to discover triggers responding in the healthy mouse gut. Testing of the Nissle 1917 library over five days following gavage (~10^7 bacteria/mouse) identified 11 strains that specifically responded to in vivo growth (Data Set A.3). Ten of these, derived from three unique promoters (ynfEFGH, torCAD, and yeaR-yoaG operons) registered a memory-on state in the majority of timepoints and all mice tested (n = 4) (Fig. 2.4D and 2.4E). Promoter response was similar during parallel analysis in the inflamed mouse gut (n = 4; see below and Fig. 2.6 for experimental details), further validating these results (Fig. A.3 and Data Set A.3).

RBS variation to adjust trigger sensitivity affects in vivo sensing capacity. Variation in sensor response based on trigger RBS was most notably observed with the ynfEFGH promoter: WT, MCD5 and MCD17 RBSs showed no response throughout the screening experiment, whereas MCD10, MCD15, and MCD23 registered as memory-on in 100%, 100% and 94% of timepoints, respectively. To validate these findings, ynfE15 (PAS816), yeaRWT (PAS817), and ynfEWT (PAS818, used here as a negative control) strains were administered to SPF mice as monocultures, with memory status determined by culture on indicator plates (at Day 2, ynfE15: 99% ± 1% SE; yeaRWT: 31% ± 4% SE; ynfEWT: 0% ± 0% SE; n = 3) (Fig. 2.4F). Of note, changing the ynfE trigger RBS from WT to MCD15 increased memory-on rate in the gut from 0% to nearly 100%.

The NGF-1 library was also screened in the healthy mouse gut over seven days following gavage (~10^7 bacteria/mouse). High library diversity was maintained, with 94% of the strains identified in the gavage sample present at the experiment endpoint (Data Set A.4). The screen
identified 22 strains that specifically responded to *in vivo* growth (Fig. A.4 and Data Set A.4), although no follow-up was conducted with individual strains.

Together these results demonstrate the ability for HTMS analysis to rapidly identify biosensors *in vivo* and the power of varying trigger sensitivity to tune the strength of biosensor response.

### 2.3.5 *In vitro* induction of *in vivo*-responding biosensor strains

The Nissle 1917 library includes some promoters with previously characterized induction conditions. We tested induction of *ynfEFGH* promoter trigger memory strains, which derive from an operon known to respond to low nitrate (through repression by phosphorylated NarL), and anaerobic (through FNR activation) conditions (115). When grown anaerobically in rich media, PAS816 (*ynfE15*), produced a memory response (13.3% ± 0.7% SE; n = 7). With added nitrate (20 mM), or in aerobic conditions, the response was zero (-O$_2$/+nitrate: 0% ± 0% SE; +O$_2$/+nitrate: 0% ± 0% SE; +O$_2$: 0% ± 0% SE; n = 3) (Fig. 2.4G), consistent with previous literature reports of *ynfEFGH* promoter behavior (115). Additionally, no response was observed from PAS818 (*ynfEWT*) in any condition (all conditions: 0% ± 0% SE; n = 3), suggesting that it is less sensitive than PAS816 (*ynfE15*), which is consistent with our *in vivo* results. *ynfE15* was also shown to respond to *in vitro* growth in cecal contents (3.7% ± 0.9% SE; n = 3) (Fig. 2.4G).

### 2.3.6 Competitive colonization experiments with operon deletions

We tested whether the native operons from which our *in vivo*-responsive biosensors are sourced provide a colonization advantage in the gut. Competitive colonization experiments were conducted by mixing a streptomycin-resistant—but otherwise wild-type—*E. coli* NGF-1 in equal amounts with a Δ*ydjLKJIHG* or Δ*ydiL-ydiM* mutant strain and administering the mixture to SPF mice. Fecal samples were collected for eight days to track the competitive index of the WT vs.
mutant strains. A null-deletion strain was used as a negative control and a ΔarcA strain was used as a positive control, based on its known growth defects (116, 117). In mice, the competitive index for the ΔydjLKJIHG and ΔydiL-ΔydiM strains remained close to that of the null-deletion control, suggesting that these operons do not confer a competitive benefit in the gut (at Day 8, ΔydjLKJIHG: 1.4 ± 0.4 SE; ΔydiL-ΔydiM: 1.6 ± 0.5 SE; Δnull: 1.4 ± 0.2 SE; ΔarcA: >6.2x10⁵, ± 4.5x10⁵ SE; n = 3) (Fig. 2.5). This is consistent with previous work showing that ΔydjLKJIHG E. coli colonized at a similar level to WT E. coli in germ free mice (84). Also notably, while there has previously been ambiguity as to the essentiality of ydiL (118, 119), our ΔydiL-ΔydiM strain grew without noticeable defects in vitro and in vivo, suggesting ydiL is not an essential gene in E. coli.

For biosensors from our Nissle 1917 library, previous studies have shown that ΔynfEFGH and ΔtorCAD mutants experience little-to-no colonization disadvantage (less than one-log difference) compared to WT strains (120). Thus, we tested a ΔyeaR-ΔyoaG mutant strain against WT E. coli NGF-1, but also found the competitive index to be similar to our null-deletion control (at Day 8, ΔyeaR-ΔyoaG: 0.7 ± 0.1 SE; n = 3) (Fig. 2.5).
Figure 2.5: Competitive index for NGF-1 wild-type strain vs. NGF-1 operon deletions. In an NGF-1 background, we deleted the native operons from which the trigger promoters in our *in vivo*-responsive biosensors were sourced, and also created positive (ΔarcA) and negative (Δnull) controls. For each operon deletion, an equal mixture of the mutant and wildtype strain was administered to BALB/c mice (ΔydiLydiM, ΔydiJKJIHG, Δnull, and ΔarcA) or C57BL/6J mice (ΔyeaRyoaG), and the competitive index was tracked by counting CFU of WT and mutant strains from fecal samples. Data points with upward pointing arrows indicate minimum competitive index on days when the mutant strain was below the limit of detection.

2.3.7 Identification of disease-specific biosensors

To look for sensors responding differentially to disease, we compared the response of the Nissle 1917 library in healthy mice (n = 4; as previously displayed in Fig. 2.4D and 2.4E, and Data Set A.3) to a murine intestinal inflammation model (Fig. 2.6A, 2.6B and A.5, and Data Set A.3). After library gavage, SPF mice were provided water containing 4% w/v dextran sulfate sodium (DSS) *ad libitum* for five days, and HTMS analysis was performed on fecal samples throughout. Weight loss (Fig. A.5A), colon length reduction at endpoint (Fig. A.5B), and increased *E. coli* CFU counts (Fig. A.5C) were all consistent with increasing inflammation throughout the experiment. Six strains registered memory-on at more timepoints in the DSS-treated group than in the control group (Fig. 2.6A and 2.6B). In particular, the ynfE17 trigger strain (PAS819)
responded specifically in DSS-treated mice (control: no response; DSS-treated: 93% of timepoints with odds ratio ≥ 0.02 and p < 0.05) (Fig. 2.6A and 2.6B, and Data Set A.3).

Figure 2.6: Library screening and individual sensor testing identifies biosensors with increased response during inflammation. (A) Screen of Nissle 1917 library in DSS-treated C57BL/6J mice (n = 4) over five days and comparison with results of previous screen in healthy C57BL/6J mice (n = 4; control group; also see Fig. 2.4D and 2.4E). To identify sensors responding more strongly to the inflamed state, the fraction of timepoints registering memory-on (odds ratio ≥ 0.02 and p < 0.05) in the control group was subtracted from the fraction in the DSS-treated group for each strain in the library, and strains were sorted according to the greatest difference between the two groups. Odds ratio heat map of the six strains registering a positive difference between the two groups is shown. Blank spaces on heat map represent insufficient sequencing coverage. See
Figure 2.6 (continued) Data Set A.3 for full results. (B) Differences in percentage of positive timepoints (odds ratio ≥ 0.02 and p < 0.05) between DSS-treated and control group mice for strains showing a positive difference between the two groups (heat map shown in Fig. 2.6A). (C) Response of HTMS strain containing the ynfE trigger with MCD17 (PAS819) in DSS-treated mice (n = 4) and healthy mice (n = 8). Memory response was assessed by plating of HTMS bacteria recovered from fecal samples on indicator plates. Error bars represent SE.

To validate ynfE17 (PAS819) response to inflammation, the strain was administered to SPF mice as a monoculture, after which a subset of the mice was provided water containing 4% DSS. Fecal samples were cultured for memory bacteria on indicator plates for 7 days following gavage. As above, body weights (Fig. A.5D), post-dissection colon lengths (Fig. A.5E) and CFU counts of HTMS bacteria (Fig. A.5F) reflected increased inflammation in DSS-treated mice. Confirming the screen results, ynfE17 showed increased response in DSS-treated mice compared to untreated controls, with the greatest difference between groups at days 6 and 7 (at Day 6, +DSS: 24% ± 9% SE, n = 4; control: 5% ± 2% SE, n = 8) (Fig. 2.6C). The strong response of PAS819 in the absence of DSS in one of the control group mice indicates that in vivo conditions other than DSS treatment can activate the ynfE17 trigger. In vitro anaerobic growth both in rich media and in cecal contents did not induce ynfE17 (0 ± 0% SE; n = 7; 0 ± 0% SE, n = 3, respectively), in contrast to ynfE15 (Fig. 2.4G), suggesting a lower nitrate threshold for ynfE17 activation and that individual bacteria experience low nitrate conditions within the inflamed mouse gut.
2.4 Discussion

Here, we have expanded the use of a robust genetic memory circuit to assess the \textit{in vivo} responses of multiple bacterial sensors in parallel. The original memory circuit (103) was altered to allow off-to-on transitions in the presence of constant induction, and to enable selection of memory-on strains from pooled libraries using spectinomycin. We developed a screening, sequencing and analysis pipeline to efficiently identify \textit{in vivo}-responding trigger–RBS combinations. Tests conducted with both genome-wide and curated libraries containing hundreds of sensors demonstrated that our method is an effective, non-invasive way to identify new biosensors responding in the gut. We identified and validated biosensors responding to growth in the healthy mouse gut and preferentially in inflamed conditions. Together, these results demonstrate the power of tuning trigger sensitivity to physiological responses and for the HTMS to assess unique features of the mammalian gut environment \textit{in vivo}.

One advantage of our method is its ability to discover sensors that could not be rationally designed based on existing knowledge, presenting an opportunity to apply the rapidly increasing but largely uncharacterized genetic diversity identified through microbiome sequencing. For instance, the two validated healthy-gut sensors from our \textit{E. coli} MG1655 library (PAS813 and PAS814) are derived from operons with largely uncharacterized function and regulation. PAS814 is triggered by the promoter of the \textit{ydjLKJIHG} operon, which putatively includes a kinase, a transporter protein, two dehydrogenases, an aldolase and an aldo-keto reductase. Only the activity of the aldo-keto reductase, YdjG, has been confirmed through reduction of methylglyoxal (121, 122). Interestingly, a previous analysis of \textit{E.coli} protein expression in germ-free mice showed that YdjG was expressed 3.5-fold higher in the cecum than \textit{in vitro} (84). Another gene
which has been studied in this operon, \textit{ydjK}, may play a role in osmotolerance, showing a $50\%$ increased growth rate in high-salt media when overexpressed in \textit{E. coli} (123). It is not known whether methylglyoxal or osmotic stress can directly trigger transcription of the \textit{ydiLKLHIG} operon. However, methylglyoxal occurs in many foods and is also produced by intestinal bacteria (124); it can also inhibit bacterial growth, suggesting a possible motivation for expression of \textit{ydiLKLHIG} in the gut.

Promoters derived from three unique \textit{Nissle 1917} operons (\textit{ynfEFGH}, \textit{torCAD}, and \textit{yeaR-yoaG}) showed memory response in the healthy mouse gut (Fig. 2.4D and 2.4E). The \textit{ynfEFGH} operon encodes a DMSO reductase which has also been shown to reduce selenate (120, 125). It is activated by FNR in anaerobic conditions and repressed by phosphorylated NarL in the presence of nitrate (115), which was further confirmed by our \textit{in vitro} tests (Fig. 2.4G).

Tuning of trigger sensitivity (e.g., by RBS modulation) is important for generating responses to physiological conditions of interest and for successful application in synthetic engineered circuits. As we observed, RBS tuning can be used to increase the response of the \textit{ynfE} promoter to as high as $100\%$ in healthy mice (PAS816; Fig. 2.4F), and to adjust the response to distinguish the inflamed gut state (PAS819; Fig. 2.6C). Importantly, the sensors we identify can be used directly in downstream applications with the memory circuit. This provides an engineering advantage over any responsive genetic elements identified through analysis in their native context, for which incorporation into synthetic circuits would routinely require additional optimization.

Inflammation leads to an increase in nitric oxide produced by the host, which generates nitrate in the intestine (126). However, because the \textit{ynfE} promoter is activated by a decrease in nitrate, our results suggest that DSS-induced inflammation may lead to lower levels of free nitrate
available to *E. coli* NGF-1, possibly due to increased local competition for nitrate via respiration by NGF-1 and other *Enterobacteriaceae*. This idea is supported by our observation of increasingly higher NGF-1 bacterial loads in fecal samples of DSS-treated mice (Fig. A.5C and A.5F), suggesting a bloom of *E. coli*—and potentially other *Enterobacteriaceae* capable of nitrate respiration—correlated with increasing duration of DSS treatment. This is consistent with previous descriptions of *E. coli* experiencing a growth advantage due to anaerobic respiration of host-derived nitrate (126). Thus, we hypothesize that PAS819 responds in DSS-treated mice specifically through sensing inflammation-induced changes in its own microenvironment.

The HTMS enables both the recording of transient signals and the amplification of low-abundance signals through antibiotic selection. These features serve as a useful complement to other techniques, such as meta-transcriptomic or -proteomic studies which capture an instantaneous snapshot of total RNA or protein content. Our use of the *E. coli* NGF-1 strain as a chassis allows reliable colonization of the mouse gut without requiring antibiotic maintenance, leading to retention of high bacterial loads and high library complexity in fecal samples even after long periods in the gut.

In creating new libraries or examining new disease models, the choice of library sources affects the potential outcome of the screen. Screening of genome-wide libraries increases the chances of discovery of new, uncharacterized systems, while curating libraries to focus on a subset of sensors can allow greater fine-tuning and increase the chances of identifying a response for the condition of interest. In addition to the promoter–RBS combinations described in this study, future libraries—especially those sourcing triggers from heterologous strains or even species—may benefit from triggers incorporating entire operons, which may increase the likelihood that any given trigger promoter has the infrastructure it requires to function.
An expanded arsenal of characterized sensors presents opportunities to construct more complex disease-responsive circuits. For instance, the combination of multiple redundant sensors would increase response accuracy under variable *in vivo* conditions, while complementary sensors may allow “fingerprinting” of different disease states. An exciting possibility is the use of more complex logic and signal processing within a single engineered strain, which may sense multiple inputs and produce anti-inflammatory, antimicrobial or other therapeutic proteins only when a precise set of conditions is satisfied (44). Sensors responding differentially based on localization within the intestine may create opportunities for more targeted drug delivery or for the construction of new safety and containment mechanisms—another important consideration in the deployment of engineered organisms.

In addition to *E. coli*, the hundreds of bacterial species naturally found in the mammalian gut encompass vast genomic diversity. Our methodology can be applied to other bacterial strains and disease models to identify systems that are capable of responding to environmental factors and biomolecules produced by the host or other microbiota. While the memory circuit we use here was selected for its demonstrated robustness in gut *E. coli*, our library construction, screening, sequencing and analysis pipeline may also be applied using other repressor- or recombinase-based memory systems in other bacterial species.

The potential to engineer synthetic circuits into commensal gut bacteria is a promising new approach to the management of intestinal disease. Synthetic biology is just beginning to tap into the evolutionary breadth of capabilities found in natural systems, and our method represents a practical means for expanding the toolkit of useful sensors for *in vivo* application.
2.5 Materials and methods

2.5.1 Media and culture conditions

Unless otherwise mentioned, bacterial cultures were grown at 37°C in LB broth or agar (10 g/L NaCl, 5 g/L yeast extract, 10 g/L tryptone). Liquid cultures were shaken at 225 rpm. Mixed liquid cultures (i.e., libraries) were grown in LBPS, which contains Peptone Special (Sigma) instead of tryptone. To quantify memory response on indicator plates, agar was supplemented with 60 µg/ml X-gal.

2.5.2 High-throughput memory system (HTMS) construction

The spectinomycin resistance gene, \( aadA \), was added downstream of the \( P_L \) promoter in the original memory switch (103) by overlap extension PCR to create the HTMS. The HTMS was genomically integrated by \( \lambda \) Red recombineering (127) into \( E. coli \) TB10 (128) between \( mhpR \) and \( lacZ \), enabling the use of endogenous \( lacZ \) as a memory-on reporter. From TB10, the HTMS was transduced by P1 \textit{vir} (129, 130) into streptomycin-resistant \( E. coli \) NGF-1 to create PAS811, which served as the chassis for biosensor strain and library construction.

2.5.3 Biosensor strain and library construction

All triggers were cloned into pDR07 (Fig. A.2), a Tn7 transposon insertion plasmid derived from pGRG36 (110). pDR07 contains a promoter-less \( cIDN \) gene and chloramphenicol-resistance cassette, flanked by Tn7 attachment sites. Two BsaI sites directly in front of the \( cIDN \) start codon allow for scarless insertion of promoter–RBS sequences into the trigger via Golden Gate assembly (12) (Fig. 2.2A). Assembled trigger plasmids were electroporated into PAS811. Transformed PAS811 cultures were grown in SOC at 30°C for 90 minutes then resuspended in LB with ampicillin (100 µg/ml) and grown at 30°C overnight to select transformants. Cultures
were then back-diluted 1:100 into LB containing chloramphenicol (25 µg/ml) and 0.1% arabinose to induce transposase genes. After 6 hours at 30°C, cultures were back-diluted 1:100 into LB with chloramphenicol and grown at 42°C for at least 6 hours to cure the temperature-sensitive plasmid while selecting for integration. The back-dilution and curing were repeated a second time.

The above steps were done on an individual or pooled basis. For individual strains, post-cure cultures were plated on agar containing chloramphenicol and individual colonies were checked for integration at the attTn7 site by PCR and Sanger sequencing. Plasmid loss was confirmed by restreaking on agar containing ampicillin. For pooled libraries, the twice-cured cultures were stored as frozen stocks and also plated on agar containing ampicillin to confirm plasmid loss. Library composition was confirmed by PCR of the trigger regions and sequencing by Illumina MiSeq.

### 2.5.4 Assessment of memory state by LacZ assay

Cultures or fecal supernatants containing memory bacteria were plated on agar plates containing streptomycin (200 µg/ml), chloramphenicol (34 µg/ml), and X-gal (60 µg/ml). The percentage of memory-on colonies was assessed by counting blue (on) and white (off) colonies.

### 2.5.5 In vitro induction

Overnight liquid cultures were back-diluted 1:100 into fresh media containing inducer, before 4 hours growth and plating on X-gal agar. For induction in cecal contents, contents of ceca from three female SPF C57BL/6J mice and suspended at 10% w/v in phospho-buffered saline (PBS). Suspensions were vortexed 90 sec and centrifuged for 3 minutes at 4300 rcf. The supernatant was recovered, supplemented with 200 µg/ml streptomycin and used for growth of HTMS bacteria.
For anaerobic inductions, pre-reduced anaerobic medium was used, and growth occurred in an anaerobic chamber (Coy Laboratory Products) under 7% H₂, 20% CO₂, 73% N₂.

2.5.6 *In vivo* induction of strains and libraries

The Harvard Medical School Animal Care and Use Committee approved all animal protocols. Experiments were conducted in female 7- to 14-week-old BALB/c (Charles River; MG1655 library) or C57BL/6J (Jackson; Nissle 1917 library) mice. Before experiments, all mice were confirmed to be absent of native streptomycin- and chloramphenicol-resistant flora. Food and water were removed ~4 hours before each gavage; water was replaced immediately, and food was replaced < 2 hours following gavage.

One day prior to bacterial gavage, mice were provided streptomycin (20 mg in PBS) by oral gavage. The next day, overnight cultures of memory strains or libraries were washed once, then diluted 10-fold in PBS and administered by gavage (100 µl; ~10⁷ bacteria/mouse).

Gavage suspension and fecal samples were plated to track bacterial load and, for individual strains, to assess memory state. Libraries were processed according to the post-exposure processing protocol below. To plate fecal bacteria, fecal samples were suspended at 100 mg/ml in PBS, vortexed 5 minutes, and centrifuged 20 minutes at 4 rcf to obtain fecal supernatant.

For inflammation experiments, water containing 4% DSS (36,000-50,000 M.W.; MP Biomedicals 160110) was substituted 2 hours after bacteria administration. Mice were dissected at the end of the experiment to measure colon length.

2.5.7. Post-exposure library processing

Fecal supernatant or *in vitro* culture was diluted 1:100 into LBPS-chloramphenicol (25 µg/ml) to achieve ~10⁶ CFU/ml. Concurrently, an overnight culture of the positive normalization
strain (PAS812 for the MG1655 library and Nissle 1917 library; ADN149 for the NGF-1 library) was back-diluted 1:100 into LBPS-chloramphenicol. Cultures were grown 4 hours, or until OD$_{600}$ ~1. The OD$_{600}$ of the positive normalization strain was adjusted to match the library culture, then diluted 1:1000 into the library culture. The resulting mix was back-diluted 1:1000 into 50 ml LBPS-chloramphenicol, and immediately split into two 25 ml volumes. Spectinomycin (50 µg/ml) was added to one culture, and both were grown overnight before centrifugation to collect bacterial pellets, which were stored at -80°C.

**2.5.8 Library sequencing and odds ratio calculation**

Genomic DNA was extracted from frozen cell pellets using a Qiagen DNEasy Blood & Tissue Kit. Using genomic DNA as a template, trigger regions from HTMS libraries were amplified by PCR and sheared with a Covaris M220 ultrasonicator to 200-600 bp fragments. Sheared products were prepared using the New England Biolabs NEBNext Ultra II Prep Kit and sequenced by Illumina MiSeq.

Raw reads were trimmed using Trimmomatic 0.36 (131) and aligned to a reference file (Data Sets A.5, A.6 and A.7 for MG1655, Nissle 1917, and NGF-1 libraries, respectively) using BWA mem 0.7.8 (132). The number of uniquely-mapped reads for each trigger was counted.

Odds ratio is expressed as $(T_x\_spect/PNS\_spect)/(T_x/PNS)$, where $T_x$ and $T_x\_spect$ are the number of mapped reads for a particular trigger in the untreated and spectinomycin-treated culture, respectively, and PNS and PNS$_{spect}$ are the number of mapped reads for the positive normalization strain in the untreated and spectinomycin-treated culture, respectively. Triggers with < 5 reads in the gavage suspension were discarded, unless they registered > 20 reads at any subsequent timepoint. For each pair of untreated and spectinomycin-treated cultures (from a
single fecal sample), odds ratios were calculated for each trigger with \( \geq 5 \) reads in the untreated culture. Statistical significance was assessed with a one-tailed Fisher’s exact test \((H_0: \text{odds ratio} = 0.02; H_a: \text{odds ratio} > 0.02)\).

For each screening result, strains with odds ratio \( \geq 0.02 \) in the gavage were discarded, and the remaining strains were sorted according to the percentage of time points that registered odds ratio \( \geq 0.02 \) and \( p < 0.05 \) over the course of the experiment.

The odds ratio calculation compares each trigger only with itself (between spectinomycin-treated and untreated cultures), normalizing any sequencing length bias between triggers. It also normalizes to the positive normalization strain in each sample, negating read depth disparities between samples.

The ability of our sequence alignment pipeline to effectively differentiate between multiple RBS variants of the same promoter was also confirmed through \textit{in silico} generation of fastq files and testing the alignment of these fastq files using our pipeline. Fastq files were generated using Grinder (133) based on a series of random and assigned distributions. The generated fastq files were designed to simulate shearing length, read lengths and sequencing depth used in our study, as well as a typical Illumina error profile. For sequencing protocols involving shearing of trigger region amplicons, we confirmed that the ratio of aligned reads to actual reads present in each simulated fastq file remained consistent for each strain in the library, regardless of the distribution used to generate the fastq file.

\textbf{2.5.9 Operon deletions and competitive colonization experiments}

Operon deletions were created in a streptomycin-resistant \textit{E. coli} NGF-1 background by \( \lambda \) Red recombineering, replacing the coding regions in each target operon with a kanamycin
resistance cassette (127). For competitive colonization experiments SPF mice were gavaged with an equal mixture of the streptomycin-resistant wild-type strain and a deletion strain. The gavage was administered, fecal samples collected, and fecal supernatant obtained as described above (see section 2.5.6). Fecal supernatant was plated on agar with both kanamycin (50 µg/ml) and streptomycin (200 µg/ml), or streptomycin alone to distinguish between total CFU and NGF-1 deletion CFU, respectively.
Chapter 3: Falling-edge pulse counters record repeated discrete occurrences of a stimulus
Preface

Much of the work in this chapter is based on the initial work of Michael Certo, who constructed the GFP memory switch and the original phage λ-based two-counter mechanism (strains PAS481-484). David Riglar provided valuable input, as well as inspiration for connecting the two-counter with the repressilator circuit. Bryan Hsu and Isaac Plant generously provided comments on this chapter.

3.1 Abstract

Synthetic biology has enabled the addition of new, programmed functions to living systems. A variety of synthetic circuits have been constructed to sense and record information, with potential applications in industry and medicine. Pulse counters are a specific type of recording circuit that can track multiple distinct exposures to a single, specific stimulus. An ideal cellular pulse counter should have a high signal-to-noise ratio and a readout that is interpretable at a single-cell level. It should also record exactly one count for each discrete stimulus pulse. However, existing counter circuits—which respond to the rising edge of a stimulus pulse and lack autoregulation—function only within a narrow range of pulse durations and concentrations. Here, we construct pulse counters that respond to the falling edge of a stimulus pulse, based on components from lambdoid bacteriophages. We use our counters to sense the presence of antibiotics in the mammalian gut and also take steps to integrate them with a synthetic oscillator circuit. Our counter shows robust counting behavior and could be expanded to count larger numbers of pulses or to accommodate different inputs. Our design provides a framework for
constructing circuits that reliably monitor, record and respond to recurring extracellular and intracellular signals.

3.2 Introduction

Synthetic biology seeks to adapt and repurpose natural systems to create new, living cells with programmable functions. Natural and engineered parts, arranged to form synthetic circuits, can allow cells to interact with their environment in increasingly complex ways. Such circuits have potential applications in biosensing, biomanufacturing, and the diagnosis and treatment of disease (6). One of the earliest benchmarks for synthetic biology has been the extent to which it can emulate the capabilities of electrical circuits, including signal transduction, processing and information storage. Toward this goal, a variety of biological circuits have been developed to carry out logic operations (8) and allow various methods of signal recording (134, 135).

Recording circuits can be divided into analog and digital varieties. Analog recorders record the order, duration and magnitude of applied stimuli, while digital recorders record only the presence or absence of signals, based on certain thresholds. In the context of synthetic biology, most analog recorders are designed to accumulate genetic modifications in a population of cells at a rate that is proportional to the concentration of one or more applied stimuli (97–101) (see section 1.4 for a discussion of synthetic recording circuits). While analog recorders can create a running temporal record with a wide dynamic range, most only work well on a population-wide scale due to relatively high levels of noise and inefficient recording at the single-cell level. Additionally, it can be difficult to produce a distinct cellular response without applying thresholding and converting the signal to a digital component. Digital recorders offer a more
efficient way to store information and generally work more reliably at a single-cell level.

Transcription-based digital recorders such as the toggle switch (102) and phage λ memory element (103), as well as DNA-based recorders based on multi-bit recombinase arrays (104) and state machines (105), can record the presence and even the order of applied stimuli. However, these types of recorders are less useful for storing temporal information. One specific shortcoming of digital recording circuits constructed to date is that they have been poorly suited to recording repeated applications of the same stimulus.

A pulse counter is a specific type of digital recorder which tracks the number of discrete pulses of a single stimulus. Biological pulse counters might monitor extracellular factors, such as fluctuations in antibiotic concentration at remote infection sites or in biofilms, or the rate of intracellular processes, such as protein fluctuations due to bacterial division or metabolism. Pulse counters may also be used to program a cellular response after a specific number of exposures. For instance, a kill switch designed to produce a lethal toxin when a cell exits its designated “safe” operating environment (73) might use a counting mechanism to prevent activation of the toxin prior to its initial introduction into that environment.

An effective cellular pulse counter demands three key characteristics. First, the circuit must exhibit a high signal-to-noise ratio, preventing advancement of the counter in the absence of a stimulus. Second, the circuit must produce a clear digital readout, such that the count can be interpreted on a single-cell basis. Finally, the circuit must advance the count only once for each discrete stimulus pulse. This last feature has been especially challenging for previous synthetic biological counters, as multiple counts may be recorded if the duration and concentration of the stimulus pulse are not carefully controlled (106). Even with a digital recording mechanism, a sustained pulse without negative feedback can cause a counter to advance multiple times. A
distinct stimulus pulse can be defined by three phases: an increase in stimulus concentration (rising edge), a sustained period of stimulus application at or above the threshold of detection, and a subsequent decrease in concentration below the threshold (falling edge). Since the first attempts at constructing pulse counter circuits (106), several mechanisms have been proposed for how a counter might avoid erroneously registering multiple counts for a single stimulus by responding only to the falling edge of a stimulus pulse (108, 136).

Here we construct falling-edge pulse counters in *Escherichia coli* based on operators sourced from bacteriophages. Our counters consist of two parts: a trigger and an actuator. The trigger contains a promoter that responds to the presence of a stimulus and activates the actuator. The actuator is a bistable switch, which, when activated, produces a recombinase that modifies the trigger, resulting in a different trigger response on next stimulus application. In our counters, the bistability and autoregulation of lambdoid phage switches serve as a mechanism to both prevent counting in the absence of a stimulus and to ensure the count does not advance multiple times for a single pulse. Our circuits allow accurate recording of the number of applied stimulus pulses, even for pulses of high concentration and extended duration. We construct two-counter circuits based on phage λ and phage 434 parts, and further combine these parts to create an improved two-counter circuit with memory function. We deploy this two-counter memory circuit *in vivo* to record bacterial exposure to an antibiotic during transit through the mammalian gut. We also integrate our two-counter with a synthetic oscillator, which can be used to track bacterial divisions, with the goal of augmenting the range of the oscillator’s tracking ability. While the prototype oscillator–counter circuit described herein is only partly functional, we suggest simple adjustments through which the full specifications of its design may be achieved.
3.3 Results

3.3.1 The cro-triggered memory circuit assumes a sustained intermediate state in the presence of activating stimulus

We have previously constructed a bacterial memory circuit in *E. coli* which is capable of sensing and recording signals in the mammalian gut (56, 103). This circuit consists of a trigger, which responds to a stimulus, and a bistable memory switch, which is turned on by activation of the trigger, resulting in continued production of a reporter gene. The memory switch is based on the bacteriophage λ lysis–lysogeny switch, contained in the λ OR operator. This operator consists of two promoters, $P_{RM}$ and $P_r$, which control the transcription of—and, in turn, are controlled by—two repressor proteins, CI and Cro (137, 138), corresponding to the off and on states of our memory circuit (see Fig. A.1A and section 2.3.1 for description).

The production of CI and Cro is regulated by three repressor binding sites in the OR operator: OR1, OR2, and OR3 (Fig. B.1A). Repressor binding at OR3 inhibits transcription from $P_{RM}$, while repressor binding at OR1 and OR2 inhibits transcription from $P_r$. Cro, under the control of $P_r$, preferentially binds to OR3, blocking $cI$ transcription. CI, under the control of $P_{RM}$, preferentially binds to OR1 and OR2, blocking $cro$ transcription (139). Furthermore, both CI and Cro have evolved autoregulation to improve the efficiency of lysis–lysogeny switching and reduce the burden of the phage on the host. At higher concentrations, a built-in negative feedback mechanism causes each protein to repress its own production (CI will bind to OR3, and Cro will bind to OR2 and OR1) (139–141).
The original memory circuit (103) relies on the exogenous expression of Cro from the trigger to flip the memory switch from a CI (memory-off) to a Cro (memory-on) state (Fig. A.1A). Activation of this cro trigger has two effects. First, Cro produced from the trigger represses $P_{RM}$ in the memory switch by binding to OR3, which halts $cI$ transcription (142); as the remaining CI is diluted out through cell division, $P_{R}$ is no longer repressed by CI. However, sufficiently high trigger activation produces high levels of Cro, which also leads to repression of $P_{R}$. This places the memory switch into a sustained intermediate state, with both $P_{RM}$ and $P_{R}$ repressed for as long as the trigger promoter is active. When the activating stimulus is removed, production of exogenous Cro from the trigger ceases, and falling levels of Cro in the cell lead to derepression of $P_{R}$ and switching to the memory-on state.

As described in section 2.3.1, one of the modifications we made to the original circuit was the substitution of cro in the trigger with a dominant-negative derivative of cI ($cIDN$). Unlike Cro, CIDN does not repress $P_{RM}$ or $P_{R}$. Instead, CIDN binds to wild-type (WT) CI monomers expressed in the memory-off state, creating heterodimers, which are not capable of binding to DNA. As a result, $P_{R}$ becomes derepressed, leading to a transition to the memory-on state (see section 2.3.1 for more detail). This transition occurs regardless of whether the trigger is continually activated; thus, memory switching is uninhibited during sustained high activation of the cIDN trigger.

To determine the level of induction required for a cro-triggered memory circuit to assume the sustained intermediate state, we created $P_{tet}$-cro trigger (PAS483, Fig. 3.1A) and $P_{tet}$-cIDN trigger (PAS484, Fig. B.1B) versions of our memory circuit which express a green fluorescent protein (GFP) reporter in the memory-on state, integrated into E. coli K-12 MG1655 (see Table 3.1 for list of strains used in this chapter). The GFP memory circuit allows a more
facile quantification of the memory state of cultured cells via flow cytometry than the β-galactosidase (LacZ) reporter used in the original memory switch. The cro (Fig. 3.1B) and cIDN (Fig. B.1C) trigger strains were induced with anhydrotetracycline (aTc) (0-32 ng/ml), either with constant induction during growth in liquid culture (“aTc constant”), or with a period of growth in the presence of aTc, followed by a wash and a period of growth without aTc (“aTc pulse”). The response of the memory circuit was quantified by GFP fluorescence. When pulsed with increasing concentrations of aTc, the cro trigger strain exhibited a full range of memory response (EC50: 0.38-0.41 ng/ml, 95% CI). However, when increasing concentrations of aTc were applied constantly without a subsequent period of aTc-free growth, the memory response peaked at 0.5 ng/ml aTc (61.0% ± 8.1% SD; n = 3) and dropped to near-zero at 2 ng/ml aTc and higher (at 2 ng/ml aTc: 0.6% ± 0.3% SD; n = 3) (Fig. 3.1B). In contrast, the cIDN trigger strain exhibited a full range of memory response across the range of applied aTc concentrations, regardless of whether aTc was pulsed or applied constantly (pulsed EC50: 0.07-0.08 ng/ml, 95% CI; constant EC50: 0.07-0.08, 95% CI) (Fig. B.1C).
TABLE 3.1 Strains used in this study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>E. coli strain</th>
<th>Trigger (P&lt;sub&gt;mut&lt;/sub&gt;)</th>
<th>Trigger Locus</th>
<th>Actuator / Memory</th>
<th>Actuator / Memory Locus</th>
<th>Plasmid</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAS483</td>
<td>K-12 MG1655</td>
<td>λcro</td>
<td>araB/C</td>
<td>λ GFP memory</td>
<td>lacZ</td>
<td>-</td>
<td>this study</td>
</tr>
<tr>
<td>PAS484</td>
<td>K-12 MG1655</td>
<td>λcIDN</td>
<td>araB/C</td>
<td>λ GFP memory</td>
<td>lacZ</td>
<td>-</td>
<td>this study</td>
</tr>
<tr>
<td>PAS481</td>
<td>K-12 MG1655</td>
<td>attB-GFP</td>
<td>araB/C</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>this study</td>
</tr>
<tr>
<td>PAS482</td>
<td>K-12 MG1655</td>
<td>attL-λcro-term.-attR-GFP</td>
<td>araB/C</td>
<td>λ actuator</td>
<td>lacZ</td>
<td>-</td>
<td>this study</td>
</tr>
<tr>
<td>PAS820</td>
<td>K-12 MG1655</td>
<td>attB-mCherry</td>
<td>attTn7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>this study</td>
</tr>
<tr>
<td>PAS821</td>
<td>K-12 MG1655</td>
<td>attL-434cro-term.-attR-GFP</td>
<td>attTn7</td>
<td>434 GFP memory</td>
<td>mhpR/lacZ</td>
<td>-</td>
<td>this study</td>
</tr>
<tr>
<td>PAS822</td>
<td>K-12 MG1655</td>
<td>attL-434cro-term.-attR-GFP</td>
<td>attTn7</td>
<td>434 actuator</td>
<td>mhpR/lacZ</td>
<td>-</td>
<td>this study</td>
</tr>
<tr>
<td>PAS823</td>
<td>NGF-1</td>
<td>attL-434cro-term.-attR-λcIDN</td>
<td>attTn7</td>
<td>434 actuator + λ lacZ memory</td>
<td>araB/C (actuator); mhpR/lacZ (memory)</td>
<td>-</td>
<td>this study</td>
</tr>
<tr>
<td>ADN146</td>
<td>MC4100</td>
<td>attL-434cro-term-attR-GFP (no TetR)</td>
<td>attTn7</td>
<td>-</td>
<td>-</td>
<td>pLPT119* pLPT145*</td>
<td>this study</td>
</tr>
<tr>
<td>ADN147</td>
<td>MC4100</td>
<td>attL-434cro-term-attR-GFP (no TetR)</td>
<td>attTn7</td>
<td>434 actuator</td>
<td>mhpR/lacZ</td>
<td>pLPT119* pLPT145*</td>
<td>this study</td>
</tr>
</tbody>
</table>

* (143)
Figure 3.1: Design and induction of GFP memory circuit. (A) GFP memory circuit (PAS483) with \( P_{tet-cro} \) trigger in memory-off and memory-on states. During trigger activation, exogenous Cro from the trigger represses both the \( P_{RM} \) and \( P_{R} \) promoters of the memory switch, creating an intermediate state. Note: constitutive promoter and RBS driving \( tetR \) are not shown. (B) Response curves of GFP memory circuit with \( P_{tet-cro} \) trigger (PAS483), induced with varying concentrations of aTc. Memory response was assessed by GFP fluorescence after induction in liquid culture. “aTc pulse” represents 4 hours of growth in the presence of aTc, followed by 4 hours of growth in the absence of aTc. “aTc constant” represents 4 hours of growth in the presence of aTc. Error bars represent SD of three biological replicates. Data courtesy of Michael Certo. (C) Response of memory circuits with \( cro \) and \( cIDN \) triggers (PAS483, \( n = 4 \); PAS484, \( n = 4 \)) to an aTc induction time course with two stimulus pulses. aTc (100 ng/ml) was subsequently applied and washed out for growth periods of at least 4 hours. A \( P_{tet} \)-GFP strain (PAS481, \( n = 7 \)) was included as a control. Error bars represent
Figure 3.1 (continued) SD. Spots of cultures on agar media (with or without aTc) are shown below at each step of the time course. Spot images consist of a GFP color overlay on a grayscale brightfield image.

The cro-trigger memory circuit assumes the intermediate state during trigger activation, regardless of its initial memory state. To illustrate this concept, the cro (PAS483) and cIDN (PAS484) trigger strains were subjected to a time course of aTc induction consisting of periods of growth in liquid culture (each at least 4 hours long), with alternating presence and absence of 100 ng/ml aTc (Fig. 3.1C). Cultures were back-diluted 1:100 between each step, and also washed prior to back-dilution from +aTc to -aTc culture to ensure complete removal of the inducer. At the end of each growth period, response was quantified via fluorescence, and cultures were also spotted on agar media (+/- aTc, as appropriate) for visualization. A Ptet-GFP strain (PAS481) was included to confirm the presence of aTc. Both cro and cIDN trigger strains were initially in the memory-off state. The cIDN trigger strain switched to a fully memory-on state with the first application of aTc and remained on through the remainder of the induction sequence (−aTc1: 0.1% ± 0.1% SD; +aTc1: 98.8% ± 1.2% SD; −aTc2: 98.4% ± 0.6% SD; +aTc2: 95.7% ± 8.4% SD; n = 4). The cro trigger strain transitioned to a memory-on state only after removal of the first aTc stimulus and returned to the intermediate state when aTc was reapplied (aTc1: 0.0% ± 0.1% SD; +aTc1: 0.1% ± 0.1% SD; −aTc2: 99.1% ± 0.9% SD; +aTc2: 0.9% ± 1.3% SD; n = 4).

As with the original memory circuit, the GFP memory element strains attained near-100% levels of induction, showing the memory circuit’s ability to function with a variety of reporter genes. In addition, the cro-triggered circuit went from over 99% memory-on to less than 1% memory-on with the second application of aTc, demonstrating the ability of the cro trigger to effectively keep the switch in the intermediate state.
3.3.2 Falling-edge pulse counter circuits

The *cro*-triggered memory circuit serves as a falling-edge sensor and recorder, due to its ability to delay transition of the memory switch to the Cro (memory-on) state until the end of a stimulus pulse. More specifically, the trigger responds to the rising edge of a pulse, while the Cro state of the memory switch responds to the falling edge. Based on this property, we used the *cro* trigger and the phage λ lysis–lysogeny switch to construct a falling-edge pulse counter circuit. We created a two-counter circuit, which produces a fluorescent response only after two distinct applications of a single stimulus. Importantly, this circuit’s ability to count is not sensitive to extended duration or concentration of each applied pulse, as long as the concentration of inducer remains sufficiently high (> ~2 ng/ml aTc; see Fig. 3.1B). The two-counter circuit achieves this by effectively responding to the falling edge of the first stimulus pulse and the rising edge of the second.

The λ two-counter circuit consists of a trigger, which senses a stimulus, and an actuator switch (Fig 3.2A). The actuator switch is similar in concept to the memory switch; however, when activated, it produces a recombinase, rather than a reporter protein. This recombinase modifies the trigger so that the trigger produces a different response with the next stimulus pulse. As with the memory circuit, the actuator begins in the CI state, with *P*RM active. CI represses *P*R, as well as *P*L (139), which controls the production of λ Xis and Int proteins. When aTc is applied, the trigger produces Cro (Fig. 3.2B), while the terminator after *cro* ensures that *gfp* is not transcribed. This places the actuator in the sustained intermediate state, with *P*RM, *P*R, and *P*L all repressed by Cro in the presence of aTc. When aTc is removed, *P*R is derepressed, and the actuator assumes the Cro state. At the same time, *P*L is derepressed, leading to the production of Xis and Int. In phage λ, the presence of higher levels of Int than Xis enables the integration of
phage DNA into the host genome. However, when Int and Xis occur in approximately equal amounts (as during induction of a lysogen), phage DNA is excised from the host genome at the attL and attR attachment sites, which revert to the original attB site. The continued presence of Xis prevents further re-integration (144, 145). In the two-counter, production of Xis and Int in the actuator’s Cro state leads to excision of the cro gene and terminator from the trigger. As a result, the trigger is primed for the second count, and a second aTc induction leads to the expression of the GFP reporter. Figure 3.2B provides a schematic of the activity and modification of the trigger during two subsequent pulses of aTc.
Figure 3.2: Falling-edge two-counter circuits based on λ and 434 phage switches. (A) λ two-counter
Figure 3.2 (continued) circuit (PAS482). Note: constitutive promoter and RBS driving tetR are not shown. (B) Schematic of trigger response and modification during two subsequent pulses of aTc. (C) Response of λ two-counter circuit (PAS482) to an aTc induction time course with two stimulus pulses. aTc (100 ng/ml) was subsequently applied and washed out for growth periods of at least 4 hours. A P_{tet}-GFP strain (PAS481) is shown as a control. Error bars represent SD of seven biological replicates. Spots of cultures on agar media (with or without aTc) are shown below at each step of the time course. Spot images consist of a GFP color overlay on a grayscale brightfield image. (D) Agarose gel depicting bands from PCR amplification of the λ two-counter trigger at the araB/araC locus. At each step of the time course, genomic DNA from a colony of PAS482 was used as a template for PCR to check for excision of the trigger. (Full trigger length: 2.5 kb; excised trigger length: 1.8 kb.) Rightmost band corresponds to a single GFP-negative colony picked from the +aTc2 plate. NTC: no-template control. 1 Kb Plus DNA Ladder (ThermoFisher) is shown. (E) Distribution of GFP fluorescence intensity in a single culture of the λ two-counter (PAS482), in which 93% of the population has successfully undergone trigger excision as a result of the first aTc stimulus. The cutoff between GFP- and GFP+ cells is indicated by the dotted line at ~100 a.u. (F) 434 two-counter circuit (PAS822). Note: constitutive promoter and RBS driving tetR are not shown. (G) Response of 434 two-counter circuit (PAS822, n = 7) to an aTc induction time course with two stimulus pulses. aTc (100 ng/ml) was subsequently applied and washed out for growth periods of at least 4 hours. A P_{tet}-mCherry strain (PAS820, n = 4) is shown as a control. Error bars represent SD. Spots of cultures on agar media (with or without aTc) are shown below at each step of the time course. Spot images consist of an mCherry color overlay on a grayscale brightfield image.

The λ two-counter strain was tested by applying a time course of sequential pulses of aTc induction and quantifying the response of the strain via GFP fluorescence (Fig. 3.2C). Each growth period was at least 4 hours long, and aTc was washed from the culture prior to each back-dilution into aTc-free media. The λ two-counter showed robust counting behavior, with two pulses required to produce GFP fluorescence. There was virtually no GFP signal during or after
the first aTc pulse, even when the pulse was applied for up to 24 hours. At the second aTc application, a high percentage of cells displayed fluorescence (-aTc1: 0.0% ± 0.1% SD; +aTc1: 0.1% ± 0.2% SD; -aTc2: 0.2% ± 0.2% SD; +aTc2: 87.8% ± 6.8% SD; n = 7). The excision of the λ two-counter trigger was confirmed by PCR amplification of the trigger region at each step of the time course. Gel electrophoresis of the PCR product showed a reduction in length consistent with excision of the region between attL and attR (full length: 2.5 kb; excised length: 1.8 kb) (Fig 3.2D). Furthermore, analysis by PCR confirmed that any GFP-negative colonies had not undergone excision and still had the original trigger.

The λ two-counter records only a single count for the initial stimulus exposure. The stability of the Cro-saturated intermediate state ensures a lack of trigger excision prior to removal of the first stimulus. Our experiments have shown that the intermediate state of the switch can be sustained even with stimulus applications of at least 24 hours.

The λ two-counter displays a high signal-to-noise ratio (Fig. 3.2E) due to the stability of the phage λ lysis-lysogeny switch in the CI state. In both the original and the GFP memory circuits, the CI protein used to maintain the memory-off state contains an ind- mutation (109) to prevent RecA-mediated cleavage, the typical mechanism of induction in the wild-type phage during an SOS response (146). Thus, the only ways through which the memory switch can exit the CI state are direct repression of Prm or inhibition of CI’s repression capability by other means (e.g., production of CIDN monomers). In our work with the memory circuit in vitro and in the mammalian gut, we have observed virtually indefinite stability of the CI (memory-off) state in the absence of induction (56, 103) (see also Chapter 2). Because the two-counter circuit is based on the same CI–Cro switch with an ind- mutation in CI, it stably maintains the CI state prior to the first application of aTc.
The bistable nature of the CI–Cro switch also produces a digital output with a clear demarcation in fluorescence between GFP-negative and GFP-positive cells (Fig. 3.2E). Prior to the second application of aTc, all cells maintained a fluorescence below ~100 a.u. At the second application of aTc, a bimodal population was observed, in which there was a marked increase in GFP fluorescence for cells that have completed trigger excision as a result of the removal of the first aTc stimulus. Because the population of cells that have experienced two pulses does not overlap with the population of cells experiencing fewer than two pulses, the readout of the count can be inferred at a single-cell level.

Building on the λ two-counter, we expanded our system to incorporate other lambdoid phage switches, as additional switches can allow the construction of more complex circuits (see section 3.3.3 below). There are a number of bacteriophages, termed “lambdoid phages,” that have similar genomic organization and function to phage λ (138). One such phage is phage 434 (147). Despite minor differences in arrangement, the CI–Cro switch of phage 434 behaves largely the same way as that of λ. However, the sequence of all three OR operator sites in phage 434 is significantly different from the OR sites in phage λ (148). The phages are also heteroimmune, with phage 434 able to infect a λ lysogen, and vice versa. The similarity in function of the two phages suggests that an analogous two-counter circuit can be constructed based on parts from phage 434. In addition, the orthogonality of λ and 434 parts suggests that both switches can be employed in the same cell to expand the function of the counter.

To construct a 434 two-counter, we created a trigger identical to that of the λ two-counter, but with 434 cro in place of λ cro, as well as an mCherry reporter in place of GFP. The 434 two-counter retained the xis/int genes and attL/attR sites from phage λ. We also created three variants of a 434 actuator, all containing a 434 CI–Cro switch (149) with an ind- mutation
at the CI cleavage site (150), but with variable placement of the \textit{xis} and \textit{int} genes. Two variants contained \textit{xis/int} driven by \textit{P}_{\text{L}}. In one of these, closest to the \textit{\lambda} actuator design, \textit{P}_{\text{L}} was positioned directly downstream of 434 \textit{cI} gene. The other contained an intermediate region between \textit{cI} and \textit{P}_{\text{L}} (more closely mirroring the organization of the wild-type phage); because the full genome of phage 434 has not been sequenced, this region was derived from the closely related phage Sf6, which has high sequence similarity (151). The third variant placed \textit{xis/int} under control of \textit{P}_{\text{R}}, downstream of 434 \textit{cro}. This third variant (Fig. 3.2F) displayed the most reliable counting behavior (data not shown for other two variants).

The 434 two-counter was tested by applying a time course of aTc induction and quantifying the response of the strain via mCherry fluorescence (Fig. 3.2G). As with the \textit{\lambda} two-counter, the 434 two-counter recorded a single count for each aTc pulse. No fluorescent signal was detected during or after the first aTc pulse, even when it was applied for up to 24 hours. At the second aTc application, a high percentage of cells displayed fluorescence (-aTc1, +aTc1, and -aTc2: 0.0% ± 0.0% SD; +aTc2: 95.9% ± 3.3% SD; \( n = 12 \)). The excision of the 434 two-counter trigger was confirmed by PCR amplification of the trigger region. Gel electrophoresis of the PCR product showed a reduction in length consistent with excision of the region between \textit{attL} and \textit{attR} (full trigger length: 4.0 kb; excised trigger length: 3.3 kb) (Fig. B.2A). As with the \textit{\lambda} two-counter, the 434 two-counter displayed a high signal-to-noise ratio (Fig. B.2B).

The fraction of fluorescent cells at the second application of aTc was significantly higher for the 434 two-counter than for the \textit{\lambda} two-counter (434: 95.9% ± 3.3% SD; \( \lambda \): 87.8% ± 6.8% SD; \( p = 0.003 \)), suggesting that either actuator switching to the Cro state or Xis/Int-driven trigger excision was more efficient in the 434 strain than in the \textit{\lambda} strain. To test for differences in efficiency of actuator switching, we constructed a 434 \textit{cro}-triggered GFP memory circuit
(PAS821), containing the 434 CI–Cro switch with a GFP reporter downstream of cro. (While PAS821 used the same trigger as the 434 two-counter, it is functionally identical to a simple $P_{tet}$-cro trigger due to the lack of an actuator in the strain.) The 434 memory circuit was subjected to a time course of aTc induction (Fig. B.2C). The fraction of memory-on cells of this circuit after removal of the first aTc pulse did not differ significantly from that of the $\lambda$ cro-triggered memory circuit (PAS483, Fig. 3.1C) (434 memory: $93.4\% \pm 6.4\%$ SD; $\lambda$ memory: $99.1\% \pm 0.9\%$ SD; $p = 0.11$). This suggests that more efficient excision, rather than more efficient switching of the actuator, is responsible for the increased response of the 434 two-counter. Thus, it is possible that placing Xis/Int under the control of the $P_R$ promoter may produce more efficient excision than placing it under control of $P_L$.

Together, these results demonstrate our ability to construct falling-edge pulse counters based on lambdoid phage switches with high stability and signal-to-noise ratio.

### 3.3.3 Two-counter memory circuit

A key motivation for constructing sensing circuits in bacteria is the potential to deploy them in inaccessible environments, such as the mammalian gut. For instance, bacteria containing a pulse counter might be used to track the number of times they encounter a signal during transit through the intestine. After exiting the gut, the bacteria may be recovered to obtain a non-invasive readout of the registered count. While the $\lambda$ and 434 two-counters produce distinct responses to two discrete pulses of a stimulus, their reporter readout requires active presence of that stimulus, as is the case with previous pulse counter circuits (106). For a deployment-and-recovery scenario, a circuit must be able to record the count and produce a sustained readout which is interpretable even when the sensor is removed from the sensing environment.
We constructed a two-counter memory circuit (PAS823), which allows both sensing and recording of two discrete stimulus pulses (Fig. 3.3A). In contrast to the previous circuits in this chapter, which were constructed in *E. coli* K-12 MG1655, this circuit was constructed in *E. coli* NGF-1. Because we planned to test this strain *in vivo*, the NGF-1 background was used due its exceptional ability to colonize the mouse gut (112). The two-counter memory circuit consists of a trigger, an actuator based on the 434 CI–Cro switch, and a memory switch based on the λ CI–Cro switch. With the first aTc application, the circuit behaves identically to the 434 two-counter, producing Xis/Int from the actuator with the falling edge of the first pulse. On the second aTc application, the trigger produces λ CIDN, which flips the λ memory switch to the Cro (memory-on) state so that it continuously produces a LacZ reporter. As with the λ and 434 two-counters, the two-counter memory circuit is designed to respond to the falling edge of the first stimulus pulse and the rising edge of the second.

The two-counter memory strain was subjected to a time course of aTc pulses, with a third aTc-free step after the second aTc pulse. The response at each step was quantified by plating cultures on X-gal indicator plates (+/− aTc, as appropriate) (Fig. 3.3B). The strain showed nearly zero activation prior to the second pulse of aTc. The second aTc application resulted in a high percentage of cells in the on state, which persisted even after aTc was removed (-aTc1: 0.0% ± 0.0% SD; +aTc1: 3.2% ± 4.1% SD; -aTc2: 0.0% ± 0.0% SD; +aTc2: 88.0% ± 6.3% SD; -aTc3: 88.7% ± 4.1% SD; n = 3). The two-counter memory circuit demonstrates our ability to incorporate complex counting and recording logic by combining multiple lambdoid phage switches in a synthetic context.
Figure 3.3: Two-counter circuit with memory. (A) Two-counter memory circuit (PAS823). Note: constitutive promoter and RBS driving tetR are not shown. (B) Response of two-counter memory circuit (PAS823), λ two-counter (PAS482) and 434 two-counter (PAS822) to an aTc induction time course with two stimulus pulses, followed by a final growth step in the absence of stimulus. aTc (100 ng/ml) was subsequently applied and washed out for growth periods of at least 4 hours. Error bars represent SD of three biological replicates. Spots of cultures on agar media containing X-gal (with or without aTc) are shown below at each step of the time course, including spots of P_tet-GFP (PAS481) and P_tet-mCherry (PAS820) strains. Spots consist of a GFP (top two rows) or mCherry (middle two rows) overlay on a grayscale brightfield image. Spots in the bottom row are a full-color brightfield image. (C) Schematic of mouse experiment with PAS823. (D) Fraction of PAS823 cells that have experienced 1 pulse or 2 pulses of aTc during transit through the mouse gut. Yellow background
Figure 3.3 (continued) indicates ad libitum administration of aTc in drinking water. Data bars indicate mean and error bars indicate range of two replicates.

3.3.4 *In vivo* sensing of aTc in the mouse gut with the two-counter memory circuit

We conducted a pilot experiment to assess the ability of PAS823 to record multiple aTc exposures within the gut (see Fig. 3.3C for experiment timeline). After a single dose of streptomycin, PAS823 was administered to specific-pathogen free (SPF) mice (n = 2) by oral gavage (~10^7 bacteria/mouse). Fecal samples were collected over five days and plated on indicator plates to check memory state. Drinking water containing aTc (100 µg/ml) and sucrose (5% w/v) was substituted after sample collection on days 1 and 4 and removed after sample collection on days 2 and 5. On days 2 through 5, samples were plated with and without aTc, enabling the calculation of the fraction of cells that had experienced zero, one or two pulses of aTc during transit through the gut (Fig. 3.3D).

All bacteria were in the memory-off state prior to gavage and one day after administration (Fig. 3.3D). At day 2—prior to removal of the first dose of aTc—a small fraction of the cells (mean: 6%) had registered two pulses; this was consistent with observations in our *in vitro* induction (see Fig. 3.3B). An average of 29% of the cells had registered one pulse. At day 3, 24 hours after removal of the first aTc dose, an average of 48% of cells had registered two pulses and an average of 28% had registered one pulse. The increasing recorded count during and after the first period of aTc administration at days 2 and 3 may be a result of fluctuating aTc concentration in the gut due to *ad libitum* administration, as well as residual amounts of aTc in the gut in the first 24 hours after removal of aTc water.
However, at day 4, there was a further increase in cells that had registered two aTc pulses (mean: 92.5%). This suggests that a fraction of the bacteria encountered at least one additional pulse of aTc between day 3 and day 4—over 24 hours after aTc was removed from the water (Fig 3.3D). Upon administration of the second aTc dose between day 4 and 5, nearly all the cells (mean: 97.5%) had registered two pulses.

Studies have shown that gastrointestinal transit time in mice is on the order of hours (152, 153). Additionally, a study of intravenous tetracycline administration in mice showed that tetracycline is undetectable in the serum after 8 hours (154). This suggests that by day 3, any residual aTc in the would have either passed through the gastrointestinal tract unabsorbed or been absorbed and excreted via the renal or biliary route. One potential explanation for the continued aTc pulses experienced by the bacteria between day 3 and 4 is the reintroduction of aTc into the gut through coprophagy (155). Transferring mice to clean bedding each day or using wire mesh cages may mitigate this factor.

3.3.5 Integration of two-counter circuit to augment cell division tracking capabilities of a synthetic oscillator

The repressilator is one of the first rationally designed synthetic circuits. The circuit consists of three repressors, which inhibit one another’s expression in turn, producing an oscillation (156). Expression of one of the repressors (TetR) controls a fluorescent reporter; thus, the oscillations can be observed through pulses of fluorescence. Recent improvements to the original repressilator have produced a highly consistent oscillation, with repressilator cells capable of remaining in-phase across a population for hundreds of bacterial generations, in the absence of any intercellular signaling (143). Furthermore, the improved repressilator’s oscillation is closely correlated with the number of divisions a cell has undergone, regardless of cell growth.
rate. This synchrony has allowed us to use the repressilator’s oscillation to track bacterial divisions on a single-cell basis (157) (preprint). This is accomplished by taking a synchronized culture of repressilator cells, all in the same phase of oscillation, and orally administering them to mice. After transit through the gut, the phase of each recovered cell can be used to estimate the number of times that cell has divided (157) (preprint). Specifically, when repressilator colonies are plated on agar media, their growth creates a series of concentric rings, corresponding to pulses of the mVenus reporter, which is under control of the TetR repressor. The distance from the origin of a colony to the first ring of fluorescence corresponds to the phase of that cell’s oscillation at time of plating. By comparing cells plated from fecal samples to cells plated at the time of gavage, it is possible to calculate the number of divisions each cell has undergone (Fig. 3.4A, left side). However, due to the inability to distinguish consecutive peaks of mVenus fluorescence, this method is only capable of tracking cell divisions within a single period of the oscillator, corresponding to approximately 15 generations.
Figure 3.4: Augmented repressilator design. (A) Plating repressilator colonies to determine phase of oscillation. Left: The Potvin-Trottier et al. repressilator with mVenus signal due to repression/derepression by TetR (143, 157). Right: Augmented repressilator design, intended to produce mCherry in addition to mVenus starting with the second oscillation. Figure adapted from (157). (B) Design of augmented repressilator circuit. The dashed vertical line separates the genome-integrated counter circuit from the plasmid-based repressilator circuit.

To augment the repressilator's function, we sought to double the number of divisions it can track by using the two-counter to produce an additional fluorescent reporter on the repressilator's second cycle (Fig. 3.4A, right side and Fig. 3.4B). This would theoretically allow tracking for up to 30 generations. The circuit was designed to begin in a TetR state, achieved via culture in the presence of isopropyl β-D-1-thiogalactopyranoside (IPTG), which inhibits LacI binding (see Fig. B.3 for conceptual schematic). When IPTG is removed, the level of TetR decreases due to repression by LacI, and oscillation commences. This produces the first mVenus
signal and should also derepress the counter trigger. While the TetR level remains low, the
actuator should assume the sustained intermediate state due to activation of the trigger promoter
and production of 434 Cro. When the level TetR rises at the start of the next oscillation, the
trigger should again be repressed, and expression of Xis/Int should allow for excision. Thus,
when the level of TetR decreases on the second oscillation, the cell should produce mCherry in
addition to mVenus.

The 434 two-counter circuit was chosen for this purpose, since the use of 434 phage
elements prevents interference from λ CI used in the repressilator. Additionally, the tetR
repressor gene was removed from the trigger of the 434 two-counter so that its activation and
repression would be controlled solely by TetR produced from the repressilator. The 434 actuator
was integrated into the nonmotile, lacI strain, E. coli MC4100 (158), which is routinely used for
repressilator experiments. This strain was also transformed with repressilator and sponge
plasmids, pLPT119 and pLPT145 (143), respectively. Lastly, the trigger was integrated into the
strain in the presence of IPTG (2 mM) to create the augmented repressilator strain (ADN147).
The IPTG synchronizes a repressilator culture by maintaining a TetR-producing state. This also
represses two-counter trigger activation during assembly. A control strain was also created,
containing the repressilator plasmids and trigger, but no actuator (ADN146).

Colonies of the augmented repressilator strain (ADN147) and an actuator-free control
(ADN146) were cultured in IPTG overnight, then plated on IPTG-free plates (Fig. 3.5A). This
produced the expected mVenus oscillations but no mCherry signal in either strain. Of note, the
augmented repressilator formed smaller colonies than the actuator-free control, possibly due to
use of gentamycin in the media (see 3.6.7 Materials and Methods). The lack of mCherry signal in
the augmented repressilator suggested that either 1) TetR levels in the repressilator did not fall to
sufficiently low concentrations to enable transition of the actuator to the sustained intermediate
state or 2) TetR levels did subsequently rise high enough to repress the trigger a second time and
enable the completion of the transition from the intermediate state to the Cro state (and thus
expression of Xis/Int).

To examine this, we cultured the originally-assembled colonies of the augmented
repressilator strain and actuator-free control in aTc overnight, then plated the cultures on IPTG-
and aTc-free plates. After overnight aTc induction, we observed fluorescent rings of mCherry
corresponding to fluorescent rings of mVenus in the augmented repressilator strain, but not in
the actuator-free control (Fig. 3.5B). This suggests that using aTc to inhibit the ability of TetR to
repress the trigger results in eventual trigger excision. Therefore, it can be inferred that the
current form of the repressilator does not maintain a low enough level of TetR to allow the
actuator to transition out of the CI state (without the exogenous addition of aTc). This may be
addressed in future versions of the circuit by increasing the sensitivity of the trigger.
Figure 3.5: Augmented repressilator testing with and without exogenous aTc induction. (A) Augmented repressilator (ADN147) and actuator-free control (ADN146) colonies were grown in +IPTG culture overnight and plated. Brightfield, yellow fluorescence and red fluorescence channels are shown. (B) Augmented
Figure 3.5 (continued) repressilator (ADN147) and actuator-free control (ADN146) colonies were grown in +aTc culture overnight and plated. Brightfield, yellow fluorescence and red fluorescence channels are shown.

3.4 Discussion

In this chapter, we constructed synthetic pulse-sensing and pulse-counting circuits, based on a bistable, bacteriophage-derived switch with a built-in negative feedback mechanism. Uniquely for synthetic biological circuits, we have shown that our counters can respond to the falling edge of a stimulus pulse. While a rising-edge detector may be preferable for immediate sensing and recording of a single signal, the ability to respond to the falling edge of a pulse is critical for circuits designed to produce unique responses to multiple discrete pulses of the same stimulus.

Our circuits demonstrated robust counting behavior *in vitro*, with the λ and 434 two-counters exhibiting three distinct advantages over previous counters. First, our counters record only a single count for a single pulse—even when pulses were applied for up to 24 hours, at concentrations significantly above the threshold of sensing. Because the sustained intermediate state of the CI–Cro switch can be maintained indefinitely while the trigger is activated, we posit that there is no upper limit to the concentration and duration of an applied pulse that would cause the circuit to record more than a single count. Second, the counters have a tightly controlled off-state, ensuring no spontaneous switching, resulting in a high signal-to-noise ratio. Third, the counters are digital reporters, which enables an interpretable readout of the count on a single-cell basis. This improves upon previous pulse counters, which produced reporting states that overlap between populations experiencing varying numbers of pulses, and thus were only interpretable when sampled as a population (106). Single-cell readable counts are particularly
important for applications in which there may be a limited number of counter cells recovered from the environment, and therefore no ability to analyze population-level signals.

We have also created and tested, to our knowledge, the first single synthetic circuit combining two phage-derived bidirectional switches. By combining λ and 434 phage OR operators and genes, we created the two-counter with memory, which continuously produces a reporter after a cell has experienced the falling edge of a first pulse followed by the rising edge of a second. Like the λ and 434 two-counters, this circuit showed efficient counting in vitro. While the outcome of our pilot experiment in mice indicated that bacteria experienced fluctuations of inducer concentration in the gut—likely due to ad libitum administration of inducer in the drinking water—follow-up studies administering discrete boluses of inducer via oral gavage and daily cage changes to remove feces are likely to produce a clearer counting result. We are also constructing additional counter circuits that respond to environmental factors such as pH and temperature, which may be useful for tracking bacterial passages through the gut.

The construction and partial demonstration of the augmented repressilator suggests the potential of the counter circuit to track intracellular processes and to expand the function of other synthetic circuits. Although the current circuit required the addition of exogenous aTc to trigger the counter, increasing the sensitivity of the trigger would likely allow the circuit to function as intended. For instance, mutating the trigger promoter to decrease the strength of TetR binding or increasing the strength of the 434 cro RBS are promising approaches. Another potential solution may be to alter the repressilator system such that it produces a sufficiently low concentration of TetR at its nadir. This might be achieved by adding TetR binding sites to the sponge plasmid (pLPT145) or adding a degradation tag to the tetR gene. However, these latter approaches carry the potential for altering the properties of the repressilator’s oscillation. Placing
a reporter on the actuator’s Cro state may also serve as a useful control for assessing whether the actuator has switched on.

There are many other opportunities for expanding the functionality of our counter circuits. For instance, a two-counter memory circuit that responds to two consecutive falling edges of a pulse (rather than a falling edge followed by a rising edge) could be easily created through substitution of the λ \( cIDN \) gene in the PAS823 trigger with \( λ \text{ cro} \). Another useful feature might be the addition of a fluorescent reporter for the first count, either at the rising edge (in the trigger) the falling edge (in the Cro state of the 434 actuator).

Expanding the range of the count is of paramount interest. In general, increasing numbers of logic operations require an increasing number of orthogonal biological parts. Because many synthetic circuits rely on a limited set of regulatory elements (e.g., TetR–\( P_{tet} \), LacI–\( P_{lac} \), CI–\( P_{R} \), AraC–\( P_{BAD} \)), scaling is a continual challenge. A three-counter could be constructed based on the two-counter memory circuit (PAS823) by 1) expressing another recombinase from the Cro state of the λ switch and 2) adding corresponding target sites and an additional reporter in the trigger. (This would create a system that responds to the falling edge of the first two pulses and the rising edge of a third.) Expanding the count beyond three may be achieved through the incorporation of other orthogonal lambdoid switches, such as those from phages HK97, HK022, P22, phi80. In addition, the creation of engineered orthogonal CI repressors (159) suggests that it may also be possible to further engineer synthetic OR operators incorporating autorepression for this purpose. Even in the absence of additional bidirectional operators with the necessary autoregulation, it may be possible to construct falling-edge counters through the use of individual promoter-repressor pairs. One proposed design is for such a two-counter shown in Figure B.4. Because additional counts will also require additional recombinase systems, expanded
counters can take advantage of a growing set of characterized, phage-derived serine integrases (104, 160).

The ability to reliably respond to distinct stimulus pulses enables a range of potential applications in tracking and responding to extracellular and intracellular signals. As we have shown with our biosensor libraries in Chapter 2, the P_{tet} trigger used here can be easily substituted with a wide variety of sensors. Synthetic biology has taken advantage of natural and designed parts to construct complex circuits, which incorporate increasingly sophisticated logic. Our demonstration of falling-edge pulse counters expands the repertoire of synthetic circuit functions and provides a scalable design for higher-order sensing.

3.5 Materials and methods

3.5.1 Media and culture conditions

Unless otherwise indicated, bacterial cultures were grown at 37°C in LB broth or agar (10 g/L NaCl, 5 g/L yeast extract, 10 g/L tryptone). Liquid cultures were shaken at 225 rpm. Unless otherwise specified, strains were induced by adding 100 ng/ml aTc to liquid media or to agar plates. To quantify two-counter memory response on indicator plates, agar was supplemented with X-gal (60 µg/ml).

3.5.2 Strain construction

Actuator, memory and trigger elements were assembled using overlap extension PCR, Gibson assembly (11), or Golden Gate assembly (12). All elements were genomically integrated at the loci specified in Table 3.1. All assembly steps were conducted in the presence of appropriate antibiotics: kanamycin (50 µg/ml) for the λ actuator and memory element; gentamicin (10
µg/ml) for the 434 actuator and memory element; chloramphenicol (25 µg/ml) for the triggers; carbenicillin (100 µg/ml) for pLPT119; kanamycin (50 µg/ml) for pLPT145.

For strains PAS481-484 and PAS820-822: Actuator or memory elements were genomically integrated via λ Red recombineering (127) into *E. coli* TB10 (128). From *E. coli* TB10, actuator or memory elements were transduced by P1<sub>vir</sub> (129, 130) into *E. coli* K-12 MG1655. Trigger elements for PAS481-484 were genonomically integrated via λ Red recombineering into a TB10 Δ<sub>int/xis</sub> mutant strain (PAS480) constructed by Michael Certo, and subsequently transduced by P1<sub>vir</sub> into the K-12 MG1655 strain containing the appropriate actuator. Trigger elements for PAS820-822 were assembled into plasmids using a temperature-sensitive, Tn7 transposon insertion backbone derived from pGRG36 (110). From these plasmids, triggers were subsequently integrated into the K-12 MG1655 strain containing the appropriate actuator using the Tn7 transposon.

Tn7 transposon integration was conducted as follows: Trigger plasmids were electroporated into the appropriate actuator-bearing strain. Transformed cultures were grown in SOC at 30°C for 90 minutes then resuspended in LB with ampicillin (100 µg/ml) and grown at 30°C overnight to select transformants. Cultures were then back-diluted 1:100 into LB containing chloramphenicol (25 µg/ml) and 0.1% arabinose to induce transposase genes while selecting for integration of triggers (which contained a chloramphenicol resistance cassette). After 6 hours at 30°C, cultures were back-diluted 1:100 into LB with chloramphenicol and grown at 42°C for at least 6 hours to cure the temperature-sensitive plasmid while selecting for integration. The back-dilution and curing were repeated a second time. Post-cure cultures were plated on agar containing chloramphenicol, and individual colonies were checked for integration at the attTn7.
site by PCR and Sanger sequencing. Plasmid loss was confirmed by restreaking on agar containing ampicillin.

For strain PAS823: The 434 actuator element was integrated into a TB10 strain via $\lambda$ Red and subsequently P1$\text{vir}$ transduced into PAS811, an *E. coli* NGF-1 strain that contains a $\lambda$ memory element (see Chapter 2). The trigger element was then integrated into this strain using the Tn7 transposon as described above.

For strain ADN146 and ADN147: The 434 actuator element was integrated into a TB10 strain via $\lambda$ Red and subsequently P1$\text{vir}$ transduced into *E. coli* MC4100. This strain, and a wild-type MC4100 strain (as a control) were transformed with repressilator and sponge plasmids pLPT119 and pLPT145 (143). Separately, the trigger element was integrated into a K-12 MG1655 strain using the Tn7 transposon. From this strain, the trigger was transferred by P1$\text{vir}$ into the actuator-containing and wild-type MC4100 strains (both also containing the repressilator and sponge plasmids). This P1$\text{vir}$ transduction was conducted in the constant presence of 2 mM IPTG to keep the repressilator in a TetR state, ensuring repression of the trigger during assembly.

### 3.5.3 *In vitro* induction

For induction time courses, strains were grown in liquid culture for periods of at least 4 hours at each step in the time course. For +aTc cultures, 100 ng/ml aTc was added. Cultures were back diluted 1:100 between each step, with an extra wash prior to back dilution from +aTc to -aTc cultures. For some replicates, each step was as long as 24 hours. At the end of each step in the time course, 2 µl of each culture was spotted on an agar plate for imaging (with or without aTc inducer and X-gal indicator, as appropriate). For strains with fluorescent reporters, cultures were analyzed by diluting 1:100 in PBS and measuring GFP and mCherry fluorescence using a BD LSR II Flow Cytometer. For strains with a LacZ reporter, cultures were plated on agar plates.
containing X-gal (60 µg/ml), and the strain response was assessed by counting blue (on) and white (off) colonies. To compare the magnitude of responses between λ and 434 two-counters and between λ and 434 memory strains, statistical significance was determined via a t-test.

3.5.4 Imaging

Agar plates were imaged using a custom-built macroscope courtesy of the Kishony Lab and the Harvard Medical School Department of Systems Biology. All spot plates from induction time courses were imaged at 24 hours after plating. Consistent camera settings were used for each channel, and any contrast adjustments were applied consistently to all images in the series.

3.5.5 Trigger excision check by PCR

PCR was used to check for trigger excision as a result of counter induction. For the λ two-counter with the trigger at the arabinose locus, the primers used were 5’-CTCCTGAAAAATCTCGATAACTCAAAAATAC and 5’-GCTTATTAAAAAGCATTTCTGTAACAAAGC. For The 434 two-counter with the trigger at the attTn7 locus, the primers used were 5’-GATAACATGCACTATCGAGATG and 5’-TTACAAACATTAATAACGAGATGACAG. PCR products were run on an agarose gel with a 1 Kb Plus Ladder (ThermoFisher #10787018).

3.5.6 In vivo sensing in the mouse gut

All animal protocols were approved by The Harvard Medical School Animal Care and Use Committee. Experiments were conducted in female 7- to 14-week-old BALB/c (Charles River). Before experiments, all mice were confirmed to be absent of native streptomycin- and chloramphenicol-resistant bacteria. Food and water were removed ~4 hours before each gavage; water was replaced immediately, and food was replaced less than 2 hours following gavage.

One day prior to bacterial administration, mice were provided streptomycin (20 mg in PBS) by oral gavage. The next day, an overnight culture of PAS823 was washed once, then diluted
10-fold in PBS and administered by gavage (100 µl; ~10^7 bacteria/mouse). Drinking water containing aTc (100 µg/ml) and sucrose (5% w/v) was substituted after sample collection on days 1 and 4 and removed after sample collection on days 2 and 5.

The gavage suspension and collected fecal samples were plated to assess memory state. To plate fecal bacteria, samples were suspended at 100 mg/ml in PBS, vortexed 5 minutes, and centrifuged 20 minutes at 4 rcf to obtain fecal supernatant. Fecal supernatant was diluted and plated agar plates containing X-gal (60 µg/ml), as well as streptomycin (200 µg/ml) and chloramphenicol (34 µg/ml). At day 2 through 5, samples were plated on plates containing both X-gal and aTc (100 ng/ml) (to assess the fraction of bacteria that had experienced at least one pulse of aTc during gut transit) and X-gal alone (to assess fraction of bacteria that had experienced two aTc pulses during gut transit). The difference of these was used to calculate the fraction of bacteria that had experienced exactly one pulse.

**3.6.7 Augmented repressilator experiment**

After assembly, ADN146 and ADN147 strains were picked from plates containing 2 mM IPTG and grown overnight in liquid culture containing appropriate antibiotics (see section 3.5.2), as well as 2 mM IPTG or 100 ng/ml aTc. The next day, cultures were plated on agar and grown at 30 °C for 4 days to allow colonies to form fluorescent rings before imaging.
Chapter 4: Conclusion
4.1 *In vivo* use of engineered bacteria

The field of synthetic biology is gradually progressing toward the design of genetic circuits intended for real-world deployment. Circuits designed for clinical application have the potential to impart new diagnostic and therapeutic functions to bacterial and mammalian cells. These circuits will allow cells to sense and respond to their environment, enabling them to monitor disease or produce localized therapeutic responses in the body. As described in Chapter 1, only a handful of biologically-responsive circuits have been constructed in bacteria to date, most of which have only been tested *in vitro*. These circuits have been designed to sense compounds related to inflammation, several bacterial and fungal pathogens, and certain environmental factors. While these studies represent a promising step toward clinically-deployable disease-responsive strains, new approaches are needed to discover additional biosensors that respond to physiological or pathological signals. This dissertation presents advances in the discovery of new biosensors for the *in vivo* gut environment, which may prove useful in constructing engineered strains for diagnostic and therapeutic purposes. At the same time basic research and proof-of-principle circuits will continue to expand the catalog of fundamental biological parts that can serve as the basis for increasingly complex designs. This dissertation also describes the first viable synthetic pulse counter circuits, which count distinct stimulus pulses without requiring precise control of pulse strength and duration, and which produce a digital, single-cell-level readout. Overall, the advances presented herein represent modest, but fundamentally important steps toward the use of engineered bacteria as clinical tools.
It’s likely that many clinical challenges—even those specifically related to the
microbiome—may find solutions that do not require the use genetically engineered (GE)
bacteria. As in the case of *C. difficile* infection, many diseases stemming from dysbiosis might be
addressed by applying a specific combination of natural microbial strains or their metabolites. 
Nevertheless, bacteria that can be programmed to sense disease and respond with a therapeutic
compound may hold immense clinical potential. As with any new therapeutic modality, the
improvement to the standard of care should outweigh any associated risks. Encouragingly, early
results from ongoing clinical trials suggest that engineered bacteria can be tolerated in healthy
humans (161). Also, while the regulatory framework around the clinical use of engineered
bacteria is still taking shape, guidance issued by the FDA indicates that such approaches are
being given serious consideration (162).

In addition to implementing basic sensing, processing and responding functions, several
factors must be considered when engineering bacteria for clinical deployment. These include
ensuring the engineered strain’s survival, ensuring potency of the strain’s therapeutic activity,
and preventing unwanted escape and gene transfer (44).

Adding engineered functions that are not essential for an organism’s survival typically
makes it less competitive against wild-type counterparts and increases the pressure to disable or
mutate the added genes. To improve an engineered strain’s chances of survival inside the host,
synthetic circuits should be constructed using the minimal set of parts required for them to fulfill
their function, with the lowest necessary gene and protein copy numbers. Co-culture of
engineered and wild-type bacteria *in vitro* and *in vivo* may be useful for assessing an engineered
strain’s ability to compete. Additionally, new approaches are being developed for managing the
burden of engineered circuits such as adjusting their level of expression based on the engineered
cell’s stress response (17, 18). An additional aspect of burden is whether an engineered strain designed to produce a therapeutic compound can produce an adequate amount to be effective. Predicting this constitutes one of the greatest challenges and will require improved in vitro or preclinical models of bacterial pharmacokinetics in the host. However, the potential for bacterial drug production to be precisely localized to the site of disease may prove promising in this regard.

Another consideration is the bacterial species and strain used as the chassis for a synthetic circuit. The use of native commensal strains is generally preferred over laboratory strains, since the former are often better adapted for survival inside the host. While most circuits have been built in E. coli due to their genetic tractability, several examples of engineered lactic acid bacteria are presented in Chapter 1. There have also been examples of synthetic circuits constructed in Bacteroides species (163). Importantly, the choice of species determines the niche that the bacterium will occupy, which should be selected based on the clinical condition being addressed. For instance, E. coli is primarily found in the large intestine, but resides closer to the gut epithelium than obligate anaerobes such as Bacteroides. Lactic acid bacteria may reside higher in the intestine. However, the ability to colonize the gut is dependent on many factors, including the bacterial species and strain, the administered dose, and the native microbiota of the host. Furthermore, while colonization of the host may provide a long-term benefit and require less frequent dosing, transient occupation—and even built-in dependence on an exogenously-administered compound—may initially be more viable from a regulatory standpoint.

A related concern is the containment of synthetic genes and organisms. Escape of engineered strains from their intended environment may be reduced through the use of toxin-based kill switches (72, 73, 164), though such containment methods may fail under strong
evolutionary pressure. Other strategies to prevent escape include engineered auxotrophy, making the strain dependent on an exogenously administered compound (161, 165). While more difficult to implement, the most reliable strategies for preventing horizontal gene transfer may involve genome-scale recoding and the use of noncanonical amino acids in engineered organisms (165, 166).

4.2 Broader considerations for recombinant DNA technology

Beyond engineering and design factors, there are also societal considerations for the deployment of GE organisms. While our ability to manipulate biology with predictable results is still in its infancy, our capabilities have advanced immensely in the past century. For scientists, it is easy to focus only on the technological aspects of research, especially when day-to-day progress in the lab moves slowly. However, we should realize that the application of recombinant DNA technology may have unique implications and externalities. While many of the most publicized criticisms of genetic engineering unfortunately stem from unscientific foundations (167), there are many legitimate environmental and societal factors that should be considered in its implementation.

Because agriculture represents the widest environmental deployment of GE organisms to date, GE crops offer several interesting case studies. While the extent to which engineered traits intermix with native populations is a subject of ongoing debate (168), rapid, large-scale adoption of GE crops has resulted in changes to agricultural practices with both positive and negative consequences. For instance, the adoption of insecticide-producing Bt crops led to a general decline in the use of insecticide sprays (2, 169), likely benefitting farmworker health and the
environment. So far, widespread development of insect resistance to Bt toxin has been largely avoided due to careful management involving the planting of non-Bt crops alongside Bt varieties (2). On the other hand, while herbicide-resistant crops have greatly simplified weed management and allowed farms to reduce energy consumption through no-till methods, their adoption has often led to an increase in herbicide use (2, 169). This stems from an overreliance on herbicides—specifically glyphosate—for weed management, which leads to increasing emergence of glyphosate-resistant weeds, which in turn pushes farmers toward even heavier glyphosate use. Although new crop varieties with resistance to other herbicides have recently been approved, better management practices are needed to prevent a repeat of the same scenario with these new traits.

Furthermore, ensuring a robust and effective regulatory framework is crucial for the responsible deployment of GE organisms. A recent scenario, which can be considered a regulatory failure, was the sale of dicamba-resistant crop seeds in United States for two growing seasons, while actually applying dicamba remained illegal (170). Dicamba is one of the oldest herbicides, but its older formulations had been banned due to their tendency to drift after spraying. In anticipation of the U.S. Environmental Protection Agency’s (EPA) approval of a new, low-drift dicamba formulation, the U.S. Department of Agriculture (USDA) approved dicamba-resistant varieties of cotton and soybean, which were sold and planted in 2015 and 2016. However, the new dicamba formulation did not receive EPA approval until the 2017 growing season. In the two intervening years, the temptation for farmers to illegally apply dicamba was great, while fines for doing so were generally negligible. This led to widespread illegal use of dicamba, destroying hundreds of thousands of acres of non-resistant crops in neighboring fields. Even in 2017, when the new low-drift dicamba entered the market, the
damage being caused by dicamba drift was estimated to affect millions of acres across the
country. Because crop insurance largely does not cover pesticide-induced damage, the only
practical solution for most farmers growing non-resistant crops was to purchase the new,
resistant seed (170, 171).

This scenario demonstrates the need for more comprehensive regulation of GE
organisms. Indeed, for over 30 years, GE organisms have been regulated by the Coordinated
Framework for Regulation of Biotechnology (172). This 1986 agreement eschewed the creation of
GE-specific regulations, instead creating a scenario in which GE organisms are regulated by the
USDA only if they involve plant pests or pathogens, by the EPA only if they produce pesticides,
and by the FDA only if they are consumed. Due to the advent of new gene editing technologies
that do not require the use of pathogenic vectors such as Agrobacterium for plant engineering, an
increasing number of GE organisms—such as glyphosate-resistant turf grass—can be freely
marketed without oversight by the USDA, EPA or FDA (173, 174).

As the potential for the deployment of GE technologies spreads beyond agriculture, a
thorough examination of the potential consequences will remain of utmost importance. While
diagnostic and therapeutic applications of recombinant DNA technology will likely be subject to
stringent regulation by the FDA, the regulation of other technologies such as transgenic insects
or even gene drives remains unclear and will likely warrant an update of regulatory frameworks
(175–177). Until recently, any discussion of the ethical limits of genetic modification in humans
may have brought to mind Gattaca-like scenarios of the far-off future. The highly controversial
announcement of the births of the first germ line-edited humans last year (178), however, quickly
brought such concerns into the present.
Even for treatments that seem ethically unambiguous from a medical standpoint, ensuring that all of humanity can share in the benefits of biotechnology will be vital. Reimbursement of gene and cell therapies is one of the most pressing emerging issues in healthcare. While there is no question that an effective intellectual property system is vital to innovation, and that these therapies can drastically improve quality of life, their list prices can approach a million dollars per treatment (179). Additional cost–benefit analyses will be required and new reimbursement models may have to be designed to make these treatments more accessible to patients (180).

While all the challenges described above should be recognized, recombinant DNA technology should by no means, in itself, be feared. Though there is much we still do not know, we should embrace the tremendous opportunity it affords to further understand our natural world and to contribute to a better one. There has hardly been a more thrilling time to be working in the field of life sciences, and I am grateful and proud to have contributed to it in my thesis work. The remainder of 21st century promises to bring further advances in genetic engineering and synthetic biology, as new tools allow us to shape living systems in increasingly powerful ways. We must be clever enough to figure out how to apply them and wise enough to do so in a way that avoids doing harm while benefiting humanity.
Appendix A: Supplemental Material for Chapter 2
Figure A.1 (A) Original memory circuit (103) in memory-off and memory-on states. Note: constitutively expressed \textit{tetR} gene is not shown. (B) Response curves of \textit{P}_{tet}-\textit{cIDN} trigger original memory (PAS809) & an \textit{aadA} memory strain (PAS810), induced with varying concentrations of aTc. Memory response was assessed by plating on indicator plates after induction with aTc in liquid culture. Error bars represent SE of three biological replicates.
Figure A.2 Plasmid map of Tn7 transposon trigger integration plasmid, pDR07, containing BsaI sites upstream of cIDN gene for modular insertion of trigger promoter and RBS variants via Golden Gate assembly (see Fig. 2.2A). Tn7 left and right attachment sites are shown in pink.

Figure A.3 Heatmap for hits from Nissle 1917 library screen in healthy C57BL/6J mice (Fig. 2.4D and 2.4E), when screened in C57BL/6J mice treated with 4% w/v DSS in water. Blank spaces on heat map represent insufficient sequencing coverage. See Data Set A.3 for full heat map.
Figure A.4 Heatmap for top ten hits from NGF-1 library screen in healthy C57BL/6J mice. Blank spaces on heat map represent insufficient sequencing coverage. See Data Set A.4 for full heat map.
Figure A.5 Body weights, CFU counts, and colon lengths of mice in Nissle 1917 library screen and ynfE17 (PAS819) biosensor strain validation experiments. All error bars represent SE. (A) Percentage of starting body weight for DSS-treated (n = 4) and healthy control (n = 4) mice used in Nissle 1917 library screens. (B) Colon lengths at dissection for DSS-treated (n = 4) and healthy control (n = 4) mice used in Nissle 1917 library screens. (C) CFU of HTMS bacteria in fecal samples of DSS-treated (n = 4) and healthy control (n = 4) mice used in Nissle 1917 library screens. (D) Percentage of starting body weight for DSS-treated (n = 4) and healthy control (n = 8) mice used in PAS819 validation experiment. (E) Colon lengths at dissection for DSS-treated (n = 4) and healthy control (n = 8) mice used in PAS819 validation experiment. (F) CFU of HTMS bacteria in fecal samples of DSS-treated (n = 4) and healthy control (n = 8) mice used in PAS819 validation experiment.
DATA SET A.1 Screen of MG1655 library in 2 BALB/c mice over 24 hours (full heat map corresponding to Fig. 2.4A left side). See Fig. 2.4 for heat map legend.

DATA SET A.2 Screen of MG1655 library in 3 BALB/c mice over seven days (full heat map corresponding to Fig. 2.4A right side). See Fig. 2.4 for heat map legend.

DATA SET A.3 Screen of Nissle 1917 library in 8 C57BL/6J mice over five days, with and without DSS treatment (full heat map corresponding to Fig. 2.4D, 2.6A and A.3). See Fig. 2.4 or Fig. 2.6 for heat map legend.

DATA SET A.4 Screen of NGF-1 library in 4 C57BL/6J mice over seven days. See Fig. A.4 for heat map legend.

DATA SET A.5 Sequence reference file (fasta) for aligning MG1655 library reads.

DATA SET A.6 Sequence reference file (fasta) for aligning Nissle 1917 library reads.

DATA SET A.7 Sequence reference file (fasta) for aligning NGF-1 library reads.
Appendix B: Supplemental Material for Chapter 3
Figure B.1 (A) Phage λ OR operator showing OR1, OR2 and OR3 repressor binding sites. $P_λ$ is repressed by binding at OR1 and OR2, while $P_{RM}$ is repressed by binding at OR3. (B) GFP memory circuit with $P_{tet}$-cIDN trigger (PAS484) in memory-off and memory-on states. Note: constitutive promoter and RBS driving tetR are not shown. (C) Response curves of GFP memory circuit with $P_{tet}$-cIDN trigger (PAS484), induced with varying concentrations of aTc. Memory response was assessed by GFP fluorescence after induction in liquid culture. “aTc pulse” represents 4 hours of growth in the presence of aTc, followed by 4 hours of growth in the absence of aTc. “aTc constant” represents 4 hours of growth in the presence of aTc. Error bars represent SD of three biological replicates. Data courtesy of Michael Certo.
Figure B.2 (A) Agarose gel depicting bands from PCR amplification of the 434 two-counter trigger during a time course of aTc induction (corresponding to Fig. 3.2G). At each step of the time course, genomic DNA from two individual colonies of PAS822 was used as a template for PCR at the Tn7 locus to check for excision of the trigger. (Full trigger length: 4.0 kb; excised trigger length: 3.3 kb.) Rightmost band corresponds to a single mCherry-negative colony picked from the +aTc2 plate. 1 Kb Plus DNA Ladder (ThermoFisher) is shown. (B) Distribution of mCherry fluorescence intensity in a single culture of the 434 two-counter (PAS822), in which 97% of the population has successfully undergone trigger excision as a result of the first aTc stimulus. The cutoff between mCherry- and mCherry+ cells is indicated by the dotted line at ~100 a.u. (C) Response of 434 memory circuit (PAS821, n = 9) to an aTc induction time course with two stimulus pulses. aTc (100 ng/ml) was subsequently applied and washed out for growth periods of at least 4 hours. A Ptet-GFP strain (PAS481, n = 7) is shown for comparison as a control. Error bars represent SD.
Figure B.3 Conceptual schematic of augmented repressilator oscillation. (See Fig. 3.4B for augmented repressilator design.)

Figure B.4 Proposed design of a three-counter, based on individual promoter-repressor pairs, which produces a reporter only on the third stimulus pulse. In this design repressor1 controls promoter P1, and repressor2 controls promoter P2. rec.a, rec.β, rec.A, and rec.B indicate recombinases/integrase genes, and triangles indicate corresponding target sites. This design can be scaled through use of additional promoter–repressor pairs and recombinases/integrases. Note: constitutive promoter and RBS driving tetR are not shown.
Bibliography


4. Evich HB, Crampton L. 2019. FDA paves way for GMO salmon to hit market. POLITICO.


