UNIVERSITY OF CALIFORNIA SAN DIEGO

Gene Regulation in Synthetic Biology: Biosensing and Novel Tools for The Construction of Complex Genetic Circuits

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Biology

by

Leo Alexander Phillip Baumgart

Committee in charge:

Professor Jeff Hasty, Chair
Professor Steven Briggs
Professor Joseph Pogliano
Professor Gürol Süel
Professor Kun Zhang

2017

The Dissertation of Leo Alexander Phillip Baumgart is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2017

EPIGRAPH

For mad scientists who keep brains in jars, here's a tip:
Why not add a slice of lemon to each jar, for freshness.

Jack Handey

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Acknowledgements

Chapter 1, in part, is currently being prepared for submission for publication of the material. In addition to the dissertation author, the following individuals contributed to this work: Johnson, Ryan; Ferry, Michael; Cookson, Scott; Cookson, Natalie; Csicsery, Nick; Stasiowski, Elizabeth; Graham, Garrett; Hasty, Jeff. The dissertation author was the one of the primary authors of this material.

Chapter 2, in full, has been submitted for publication as it may appear in Nature Genetics, 2017, Baumgart, Leo; Mather, Will; Hasty, Jeff. The dissertation author was the primary author of this paper.

VITA

2017 Doctor of Philosophy, University of California San Diego
2011 Bachelor of Science, University of California San Diego

PUBLICATIONS

Baumgart, L., Mather, W., Hasty, J. Synchronized DNA Cycling Across a Bacterial Population. Nature Genetics (in press 2017).

Luke, C. S., Selimkhanov, J., **Baumgart, L.**, Cohen, S. E., Golden, S. S., Cookson, N. A., & Hasty, J. (2015). A Microfluidic Platform for Long-Term Monitoring of Algae in a Dynamic Environment. ACS synthetic biology, 5(1), 8-14.

J. Hasty, **L.A. Baumgart**, S. Cookson, M. Ferry, G. Graham, R. Huerta, R. Johnson, and L. Tsimring, Microbial microfluidic biosensor, August 25 2016, WO Patent App. PCT/US2016/017,889.

ABSTRACT OF THE DISSERTATION

Gene Regulation in Synthetic Biology: Biosensing and Novel Tools for The Construction of Complex Genetic Circuits

by

Leo Alexander Phillip Baumgart

Doctor of Philosophy in Biology

University of California, San Diego, 2017

Professor Jeff Hasty, Chair

One of the pursuits of synthetic biology is to harness gene expression by reprogramming cells to perform useful functions. As one of the most thoroughly characterized model organisms, *Escherichia coli* presents significant potential for engineering synthetic circuits. In this work I present both simple circuits designed for biosensing applications, as well as novel strategies that may allow combining such simple circuits into more complex networks. Specifically, I describe methods for both negative and positive dynamic modulation of plasmids in *E. coli*, allowing for control of gene expression without requiring specially engineered promoters. Combining individual synthetic modules will be necessary for encoding more complex circuit logic and function.

Chapter 1. A microfluidic biosensor for detection of toxins

Chapter 1.1 – Introduction to biosensors

As global freshwater supplies become impacted, water quality is an increasingly relevant and important topic. Human activity has introduced toxins, in particular heavy metals, into many natural water supplies across the globe. A prerequisite for properly managing water supplies intended for drinking and irrigation is the ability to easily and cheaply monitor toxin levels in water sources. In addition, tracking the ecological impacts of waterborne toxins requires frequent and accurate on-site measurements in remote locations. While analytical chemistry methods exist to detect a wide variety of heavy metals and other toxins, they require bulky, expensive equipment and highly trained technicians. Furthermore, while low cost chemical test strips exist for some toxic compounds, they require frequent manual sampling to be effective and may not be practical in a field operation. Both of these methods give only snapshots of the water systems quality and may miss spikes in the concentration of a toxin, especially in remote environments where testing can only be accomplished infrequently. Ideally, a low cost and easy to maintain system would be developed which can be installed locally at a water source to continuously report on water quality with minimal intervention. Increasingly, sensors relying on biological organisms are being considered for such applications. The advantages of biosensors include low cost, the ability to detect a wide variety of chemical

substances, and tolerance to varying environmental conditions[1]. Unfortunately current biosensors often suffer from low sensitivity and poor selectivity, which limit their potential[1]. Recent advances in synthetic biology have the potential to increase both the sensitivity and the selectivity of biosensors based on microbial organisms. In conjunction with genetic advances, microfluidic technology has the potential to make extremely low cost, lightweight, and small form factor biosensors a reality.

I worked together with a team of graduate students, lab technicians, and post-doctoral researchers on a DARPA-funded project titled "An Online Biosensor for the Protection of Water Supplies." Our goal was to develop a biosensor that can detect a variety of heavy metal toxins and ammonia in water. We successfully integrated high-throughput sequencing, microfluidics, and computational algorithms to develop a field-deployable biosensor, which can directly monitor an input stream of natural water and report on toxin content in real-time. The biosensor can reliably detect and distinguish 6 heavy metals and ammonium ions in water at levels relevant to drinking water safety. To develop this sensor, we initially surveyed the transcriptomic responses of both *Escherichia coli* and *Bacillus subtilis* to a variety of heavy metals. Using this RNA-Seq data in combination with previously characterized promoters from these and other species, we were able to select a minimal set of promoters found to be regulated specifically in response to each toxin. We designed a simple genetic sensor construct for each, where the toxin-responsive promoter controls the expression

of green fluorescent protein (GFP). For most constructs, we designed and tested several versions with small changes, such as a different plasmid origin of replication or ribosome binding site (RBS). This resulted in more than 125 distinct constructs, which we tested and compared in both batch and microfluidics to select a minimal set of maximally responsive and orthogonal strains. Additionally, we developed a novel microfluidic chip that allows long-term culturing of multiple *B. subtilis* and *E. coli* strains in separate traps. In our implementation each trap region contains a strain with a unique genetic sensor construct on a plasmid. Peristaltic pumps move liquid though the microfluidic device, drawing from both the water sample to be tested as well as a culture medium concentrate, which are mixed on-chip upstream of the cell traps. In order to monitor the GFP output of each trap region in real-time, we constructed an inexpensive optical system using off-the-shelf parts. The combined fluorescence output of the sensor strains creates a unique signature for each toxin that can be visualized within three hours of the toxin entering the influent water stream. My major contributions to this project were: (1) the identification of toxin-responsive promoters from literature and by transcriptomic profiling, (2) plasmid and biosensor construct design, and (3) bacterial growth media optimization.

Chapter 1.2 – Transcriptomic profiling by RNA sequencing

The organisms selected as biosensor chassis strains were *Escherichia coli* strain MG1655 and *Bacillus subtilis* strain 3610. These species are both well-studied model organisms and their genomes have been fully sequenced and many genes have been annotated. Additionally, there are extensive tools available for genetic manipulation of these species. In addition to conducting extensive literature search to identify promoters that were known to be responsive to the toxins of interest (Table 1.1), we conducted our own search by exposing bacteria to sub-lethal levels of toxins and profiling the transcriptomic responses. To accomplish this goal, cultures of exponential phase bacteria were cultured for several doublings (2.5-3 hours) in the presence of toxin. At the end of the exposure period, these cultures were stabilized using RNA Protect reagent, and centrifuged to pellet bacteria. These pellets were used to extract RNA, from which barcoded sequencing libraries were generated for high-throughput sequencing on an Illumina MiSeq instrument. To determining the appropriate level of toxin exposure for these RNA-Seq experiments, wild type strains of bacteria were grown in a 96-well plate with a dilution series of each toxin and growth rate was monitored using a plate reader taking OD_{600} readings at 5-minute intervals. In general, two toxin levels were identified for sequencing experiments: one that showed a very minimal growth defect, and a second level 10-fold higher than the first. In some cases where this 10-fold higher level was observed to cause a very

severe growth defect, the higher level of toxin was reduced to be 2 or 3-fold of the lower level. The rationale was that the levels should be high enough to induce a significant response from bacteria, but not so high that the growth rate would be reduced too severely, which may result in exposed samples no longer being comparable to negative control samples. Only *B. subtilis* was exposed to ammonium, since the *E. coli* growth medium already contained a high level of ammonium, which is required for growth. We reasoned that this high background level of ammonium would prevent any significant response to additional small amounts of ammonium. For a detailed description of the RNA-Seq experimental protocol, see section 1.6 – Protocols and methods.

Chapter 1.3 – Biosensor construct design

Biosensor constructs consist of a toxin-responsive promoter driving the expression of green fluorescent protein (sfGFP), and in some cases also include a gene coding for a transcription factor that acts on the toxin-responsive promoter. For promoters native to the chassis organisms, we found we could often rely on the native expression levels of transcription factors from the genome. A majority of toxin-responsive promoters identified by literature search, however, were native to other organisms and thus the relevant transcription factors were necessarily included on the biosensor construct. All heterologously expressed genes were codon optimized for our organisms to ensure that codon utilization

and GC content were suitable for expression in our chassis organisms. For *E. coli* biosensor constructs, we used a p15A origin plasmid, while *B. subtilis* constructs were integrated in the neutral *lacA* site on the genome. A typical sensor plasmid map is shown in Fig. 1.1. RNA-Seq analysis provides information on up- and down-regulated genes in response to toxins, but does not supply information on the promoter region responsible. For construction of the sensor plasmids, we defined the promoter region as 200 bp upstream of the start site of the gene unless more information was available. If detailed literature information was available that suggested regions >200 bp upstream were important for promoter regulation, we extended this length to include all known regulatory elements. All constructs were de novo synthesized in their entirety and sequence verified before transformation into the chassis strains.

Chapter 1.4 – Bacterial media optimization for microfluidics

Minimal medium was chosen for all experiments and for use in the biosensor device, because this offers several benefits as compared to rich medium. Minimal salts medium (M9) is a low cost defined medium, which allows consistency across experiments and avoids introducing trace amounts of heavy metals which may be found as contaminants in undefined mediums such as lysogeny broth (LB). In addition, amino acids have been found to chelate a large variety of heavy metals [2]. Thus, choice of a medium devoid of amino acids is

desirable to allow lower detection limits of heavy metals in water. The microfluidic biosensor chip used in our device is designed for on-chip mixing of a concentrated bacterial growth medium and water. This allows the device to draw water directly from the source being monitored for toxins.

When using a standard recipe for M9 minimal salts medium as the media concentrate, we found that over several hours the mixing region of the devices became clogged with a precipitate, preventing further water from flowing through the device (Fig. 1.2). We suspected that this precipitate was a result of calcium-phosphate crystals building up in the mixing regions. To overcome this limitation, we developed a modified form of M9 minimal salts medium, which we called HM9. This medium is a reformulation of M9 minimal salts medium but also incorporates some elements of a previously published HMM medium[3]. Since phosphate and calcium are required for bacterial growth, we could not eliminate these components from the medium entirely. We replaced the sodium phosphate with glycerol-2-phosphate, which can be utilized as an alternative source of phosphorus by both *E. coli* and *B. subtilis.* In the original formulation of M9, inorganic phosphate also serves as a buffer, which we replaced with MOPS buffer. In this modified medium we found that growth was improved when supplemented additionally with potassium salt and iron. For *B. subtilis*, we also supplemented the medium with zinc and manganese. With the HM9 medium that we developed we do not observe any precipitates accumulating in the mixing

region of the device. A detailed description of the media formulations and recipes can be found in section 1.6 – Protocols and methods.

Chapter 1.5 – Biosensor responses to toxins

We sought to identify a subset of the >125 sensor constructs that would allow specific sensing of the target toxins while minimizing the number of strains required. In an initial round of testing, all sensor constructs were evaluated for functionality in a high-throughput screen using a plate reader and 96-well plates. Each construct was transformed into the appropriate chassis strain and tested against a dilution series of the toxin expected to illicit a response. After discarding sensor constructs that produced a very weak or undetectable response in 96-well plates, the remaining strains were tested in microfluidic devices. We determined that *E. coli* strains alone were sufficient for specific detection of all the heavy metals of interest, and therefore decided to use *B. subtilis* solely for the detection of ammonia.

The native *E. coli* ArsR promoter was selected for sensing of arsenic in strain As7. This promoter has been previously characterized and shown to be responsive mainly to arsenic and antimony, via the action of the transcriptional repressor protein ArsR, which binds to the promoter region and inhibits transcription in the absence of these chemical inducers[4]. For lead, we use the CadC transcriptional repressor and the corresponding promoter, both found on

plasmid pI258, which is native to *Staphylococcus aureus*[5]. For sensing of mercury, we selected the well-characterized MerR transcriptional activator with the corresponding bi-directional promoter. MerR is naturally found in a variety of gram-negative bacteria and plasmids on the transposon Tn21[6]. While it is native to some species of *E. coli*, it is not naturally found in our strain MG1655. In addition, we use three promoters found on the MG1655 genome which we found were responsive to toxins without requiring overexpression of any additional transcriptional factors: PzntA[7] for cadmium and lead, PcusC[8] for copper, and PzraP[9] for lead. All the biosensor constructs described previously were placed on a medium-copy p15A origin plasmid with a spectinomycin resistance gene. Some crosstalk between metals is apparent in individual strains, however the strains respond to these metals to different degrees. By examining the combined responses of all strains, individual metals in the water supply can be uniquely identified.

In addition to these heavy-metal responsive genes, we use the PnasB promoter native to *B. subtilis* for sensing of ammonium. Although it would be simpler to use only a single chassis strain for all biosensor constructs, sensing ammonium with E. coli is not as straightforward. In order to elicit a robust response to low levels of ammonium entering the water supply, the bacteria must be fed a nitrogen source that is of a lower quality than ammonium. While *B. subtilis* can grow on nitrate as the sole nitrogen source, this is not the case for *E. coli*. To address this, we formulated a medium that replaces the high levels of

ammonium contained in the E. coli medium with nitrate. In these conditions, any ammonium entering the water supply will be sensed by the *B. subtilis* as a superior nitrogen source, resulting in repression of the PnasB promoter, which is involved in nitrate assimilation[10].

The final strains selected are summarized in Table 1.2. Responses for *E. coli* strains to various toxins are shown in Figure 1.3, and *B. subtilis* response to ammonium is shown in Figure 1.4

Chapter 1.6 – Figures and Tables

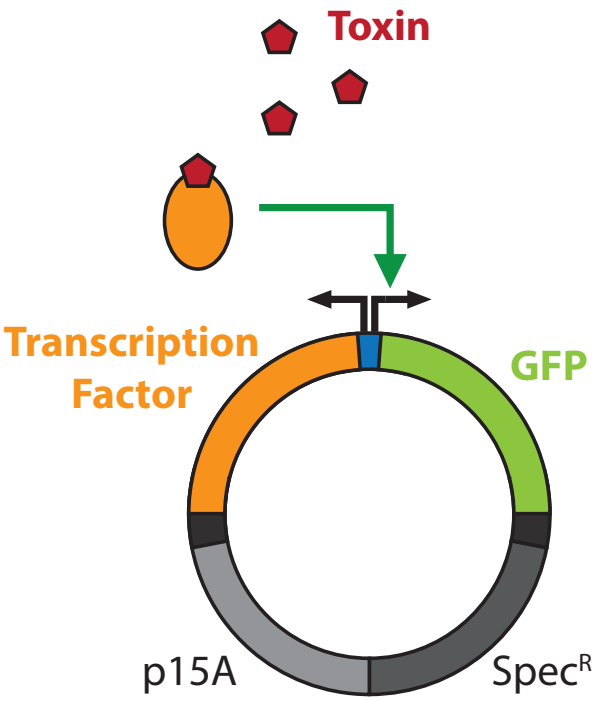


Figure 1.1. A generic sensor plasmid used for sensing of toxins in *E. coli*. The p15A origin was used for all sensor plasmids along with a spectinomycin resistance gene ($Spec^R$). The GFP used was superfolder GFP (sfGFP, shown in green). Not all plasmids contained a transcription factor (shown in orange), and in some cases the placement of promoters differs from that shown here.

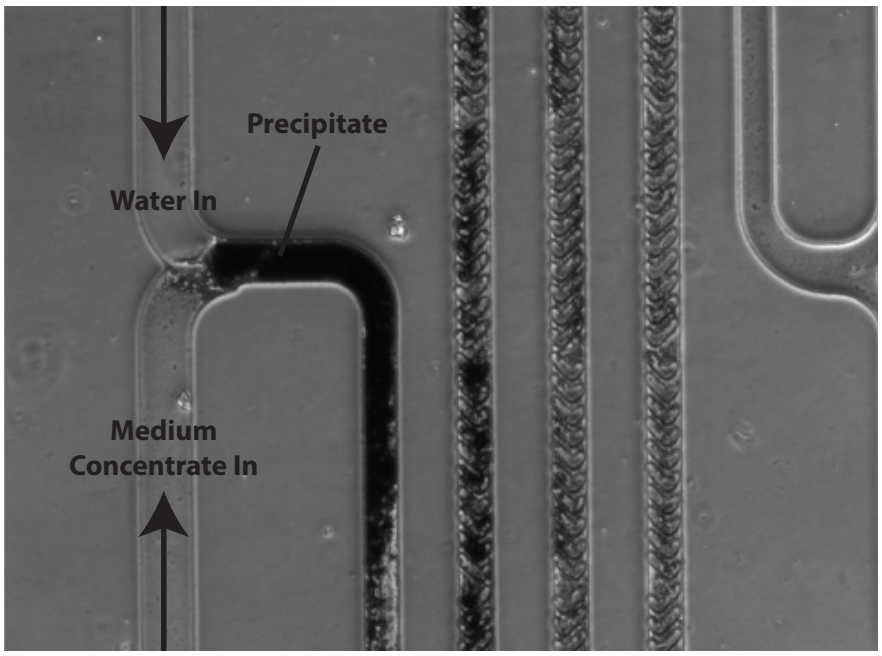


Figure 1.2. Microscope image depicting precipitate that formed when mixing concentrated M9 minimal medium with water.

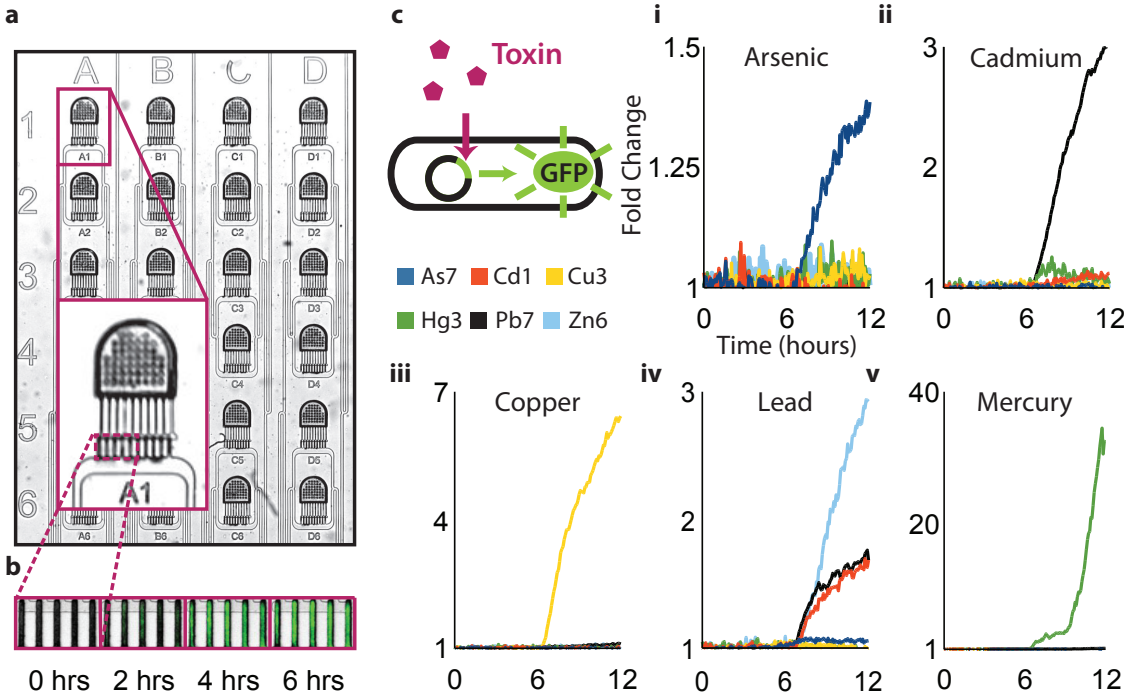


Figure 1.3. *E. coli* biosensor strain responses to heavy metal toxins. (a) Image of a microfluidic chip loaded with a different *E. coli* strain in each position. Inset in bottom left shows a single strain. (b) Higher resolution image of the region used to extract fluorescence data. Filmstrip shows induction of strain As7 in response to arsenic over the course of 6 hours. (c) When toxins enter the water supply, a fluorescent response in a unique combination of strains allows identification of the specific toxin.

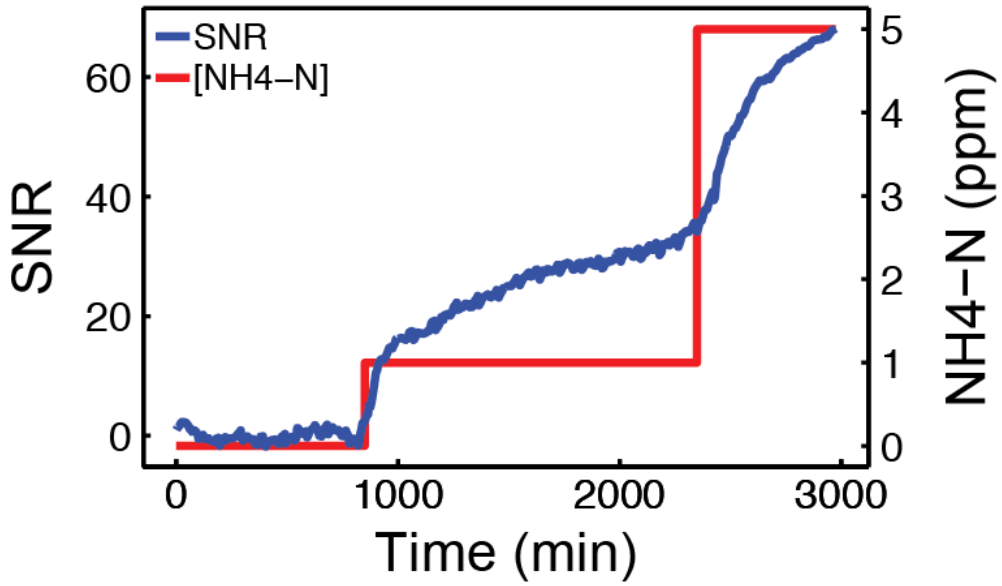


Figure 1.4. *B. subtilis* responding to ammonium. Plot shows signal-to-noise ratio (SNR) instead of raw fluorescence values, as this sensor is a "lights-off" sensor with a high degree of background noise, where fluorescence signal decreases temporarily upon induction with ammonium.

Table 1.1. Target toxins for biosensor detection and the respective chemical forms used in experiments. Concentrations listed correspond to those used in RNA-Seq experiments.

Toxin of Interest	Chemical Form	Concentrations (uM)
Arsenic	$NaAsO_2$	0.25, 1
Mercury	$Hg(NO_3)_2$	0.01, 0.1
Cadmium	$CdCl_2$	0.4, 1.2
Lead	$Pb(NO_3)_2$	0.1, 3
Chromium	$K2CrO_4$	0.2, 2
Copper	$CuSO_4$	2, 20
Ammonia	NH_4Cl	71, 710

Table 1.2. List of biosensor strains chosen for specific detection of toxins.

Host strain	Promoter	Transcription factor	Responds to	Strain name
E. coli	ParsR from R773	ArsR from R773	Arsenic, weakly to lead	As7
E. coli	PcadC from S. aureus pI258	CadC from pI258 (codon optimized)	Lead, weakly to cadmium	Cd1
E. coli	Pmer from Tn21 (bidirectional)	MerR from Tn21 (codon optimized)	Mercury, weakly to cadmium	Hg3
E. coli	PzntA from E. coli MG1655	n.a.	Cadmium, lead	Pb7
E. coli	PcusC from E. coli MG1655	n.a.	Copper	Cu3
E. coli	PzraP	n.a.	Lead	Zn6
B. subtilis	PnasB from 3610	n.a.	Ammonium	Amm3

Chapter 1.7 – Protocols and methods

Cell culturing and media

For all batch culture experiments M9 minimal medium supplemented with 0.4% w/v glucose, 0.1 mM $CaCl_2$, and 2 mM $MgSO_4$ was used. For *B. subtilis*, medium was additionally supplemented with 0.075% v/v TWEEN 20, 50 μM $FeCl_3$, 50 μM $MnCl_2$, and 1 μM $ZnCl_2$. For *B. subtilis* ammonia exposure experiments, the NH_4Cl in M9 minimal medium was replaced with $NaNO_3$, keeping the concentration of nitrogen constant. For microfluidic experiments, we developed a minimal medium optimal for growth of bacteria and heavy metal sensing, adapted from HMM[3]. This medium replaces the inorganic phosphate in M9 minimal medium with glycerol-2-phosphate, MOPS (pH = 7.2), and KCl. Inorganic phosphate is undesirable because of its metal chelation properties and propensity to form calcium phosphate deposits within microfluidic channels. To minimize contaminating metals, all microfluidic experiments were carried out with media made with extra high purity salts where available. We found that when using these pure salts, robust *E. coli* growth required supplementing the media with iron, and robust *B. subtilis* growth required iron, zinc, and manganese.

E. coli media composition for microfluidics

40 mM PharmaGrade MOPS [Sigma #PHG0007-1KG] (from 1 M stock at pH 7.2)
4 mM glycerol-2-phosphate [Sigma #G6501-25G]

0.4% w/v dextrose (glucose) [Sigma #D9434-1KG]

1 g/L (19 mM or 262 ppm NH_4-N) TraceSelect NH_4Cl [Sigma #09725-100G]

3.7 g/L (50 mM) TraceSelect KCl [Sigma #05257-100G]

0.075% v/v TWEEN 20 [Acros Organics #23336-0010]

50 µg/mL spectinomycin (as spectinomycin dihydrochloride pentahydrate) [Sigma #S4014-5G]

1 µM $FeCl_3$ [Alfa Aesar #A16231-500G]

0.01 mM $CaCl_2$ [Macron Fine Chemicals #4160-12]

0.2 mM $MgSO_4$ [Macron Fine Chemicals #6066-04]

B. subtilis media composition for microfluidics

Same as *E. coli* media with the following changes:

Replace NH_4Cl with 1.6 g/L $NaNO_3$

Use 50 µM $FeCl_3$ instead of 1 µM

add 50 µM $MnCl_2$ [Baker #2540-01]

add 1 µM $ZnCl_2$ [Macron #8780-04]

Recipes for 250 mL of 5X media concentrate for microfluidics

HM9-EC (*E. coli*)

50 mL 1 M (25X) MOPS

50 mL 100 mM (25X) glycerol-2-phosphate (G2P)

25 mL 20% w/v (50X) glucose

1.25 g TraceSelect NH_4Cl

4.63 g TraceSelect KCl

1.25 mL 75% v/v (1000X) TWEEN 20

1.25 mL 50 mg/mL (1000X) spectinomycin (as spectinomycin dihydrochloride pentahydrate)
25 µL 50 mM (50,000X) $FeCl_3$
1.25 mL 10 mM (1000X) $CaCl_2$
1.25 mL 200 mM (1000X) $MgSO_4$
Fill to 250 mL in volumetric flask, then 0.2 µm filter

HM9-BS (*B. subtilis*)

50 mL 1 M (25X) MOPS
50 mL 100 mM (25X) glycerol-2-phosphate (G2P)
25 mL 50X glucose (20%)
2 g TraceSelect NaNO3
4.63 g TraceSelect KCl
1.25 mL 75% v/v (1000X) TWEEN 20
1.25 mL 75 mg/mL (1000X) spectinomycin dihydrochloride pentahydrate
1.25 mL 50 mM (1000X) FeCl3
1.25 mL 50 mM (1000X) MnCl2
1.25 mL 1 mM (1000X) ZnCl2
1.25 mL 10 mM (1000X) CaCl2
1.25 mL 200 mM (1000X) MgSO4
Fill to 250 mL in volumetric flask, then 0.2 µm filter

Note that TraceSelect formulations of reagents were used when available to minimize the potential for heavy metal contamination of the media concentrate.

RNA-Seq

E. coli:

E. coli MG1655 was inoculated from a frozen stock (maintained at -80°C) into 3 mL of M9 minimal medium and grown overnight in a 14 mL culture tube. After overnight growth, the culture was diluted 4,000-fold into 100 mL fresh M9 medium in a 500 mL flask and allowed to grow until an OD_{600} in the range of 0.05-0.20 was reached. At this point cultures were adjusted to an OD_{600} of 0.05 using fresh media. For each RNA sample, 1.5 mL of this normalized culture was mixed with 1.5 mL of toxin-containing media in 14 mL culture tubes, resulting in a final OD_{600} of 0.025. Cells were grown with shaking at 300 rpm at 37°C for either 2.5 or 3 hours. To ensure that all RNA was extracted from exponential phase cells, cultures were not allowed to progress beyond an OD_{600} of 0.3.

B. subtilis:

B. subtilis 168 was inoculated from a frozen stock (maintained at -80°C) into 3 mL of LB medium and grown overnight in a 14 mL culture tube. After overnight growth, the culture was diluted 1,000-fold into 100 mL fresh M9 medium in a 500 mL flask and allowed to grow until an OD_{600} of 0.10-0.20 was reached. At this point cultures were adjusted to an OD_{600} of 0.10 using fresh media. For each RNA sample, 1.5 mL of this normalized culture was mixed with 1.5 mL of toxin-containing media in 14 mL culture tubes, resulting in a final OD_{600} of 0.05. Cells were grown with shaking at 300 rpm at 37°C for 2.5 hours. To

ensure that all RNA was extracted from exponential phase cells, cultures were not allowed to progress beyond an OD_{600} of 0.3. This protocol was adjusted for ammonia exposure experiments in the following ways: (1) M9-nitrate medium was used in place of M9-ammonia, (2) cells were grown to 0.20 OD_{600} before ammonia exposure, and (3) cells were pelleted for RNA extraction after 30 minutes of ammonia exposure. The short incubation time allowed for ammonia-exposed samples, which grow significantly faster than negative control samples, to be in a similar growth phase at the time of RNA extraction.

RNA Extraction

For each condition, five biological replicates were performed in parallel. Four of these replicates were used for RNA extraction and the remaining one was used for obtaining a final OD_{600} reading. For RNA stabilization, immediately after the incubation period each 3 mL culture was mixed with 6 mL of Qiagen RNAprotect Bacteria Reagent and incubated at room temperature for 5 minutes. Stabilized cells were then pelleted at 5000 x *g*, decanted to remove supernatant, and immediately stored at -80°C for two weeks or less. RNA extractions were performed using the Qiagen RNEasy kit following the manufacturer's protocol. The optional β-mercaptoethanol addition to buffer RLT was included. For *B. subtilis* only, a 10-minute digestion with proteinase K at room temperature with frequent vortexing was included prior to homogenization. For homogenization, cell pellets resuspended in 350 µL buffer RLT were transferred to 1.5 mL

Eppendorf Safe-Lock tubes prefilled with 200 µL of RNase-free 0.15 mm zirconium oxide beads. Resuspended pellets were homogenized in a Bullet Blender Storm for 5 minutes at maximum intensity. DNase digestion was performed either (1) on-column using the Qiagen RNase-Free DNase kit or (2) after RNA extraction using the Zymo RNA Clean & Concentrator-5 kit with the DNase I set following the manufacturer's protocol. From each batch of toxin exposure experiments all samples including negative controls were processed using an identical protocol.

Strains and Plasmids

The *E. coli* strain used for all experiments is wild type MG1655. A modified *B. subtilis* 168 was used for all RNA-Seq experiments. This strain was modified by knocking out the *hag* gene using Cre-Lox recombination. Additionally, tryptophan biosynthesis was restored by introducing the *trpC2* gene from *B. subtilis* NCIB 3610 (resulting genotype: *B. subtilis* 168 *hag*::*lox trpC2*+). All microfluidic experiments were conducted using a modified *B. subtilis* NCIB 3610, which was found to grow more robustly in minimal media than *B. subtilis* 168, especially when nitrate was supplied as the sole nitrogen source. This strain was modified by knocking out the *hag* and *espH* genes to reduce motility and biofilm formation (resulting genotype: *B. subtilis* NCIB 3610 *hag*::*cat epsH*::*tet*).

RNA-Seq and Illumina Library Prep

DNase treated RNA samples were enriched for mRNA with Epicentre Ribo-Zero rRNA removal kit (Gram-Negative Bacteria for *E. coli* and total Bacteria for *B. subtilis*) prior to library preparation. Libraries were generated using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina following the manufacturer's protocol. Samples were normalized, pooled into sets of up to 12, and sequenced on an Illumina MiSeq using the Illumina 150 cycle v3 reagent kit.

Chapter 1.8 – Acknowledgements

Chapter 1, in part, is currently being prepared for submission for publication of the material. In addition to the dissertation author, the following individuals contributed to this work: Johnson, Ryan; Ferry, Michael; Cookson, Scott; Cookson, Natalie; Csicsery, Nick; Stasiowski, Elizabeth; Graham, Garrett; Hasty, Jeff. The dissertation author was the one of the primary authors of this material.

Chapter 1.9 – Works cited in chapter 1

1. Su, L., Jia, W., Hou, C., & Lei, Y. (2011). Microbial biosensors: a review. Biosensors and Bioelectronics, 26(5), 1788-1799.

2. Dişbudak, A., Bektaş, S., Patır, S., Genç, Ö., & Denizli, A. (2002). Cysteine-metal affinity chromatography: determination of heavy metal adsorption properties. *Separation and Purification Technology*, *26*(2), 273-281.

3. Ivask, A., Rõlova, T., & Kahru, A. (2009). A suite of recombinant luminescent bacterial strains for the quantification of bioavailable heavy metals and toxicity testing. *BMC biotechnology*, *9*(1), 41.

4. Diorio, C., Cai, J., Marmor, J., Shinder, R., & DuBow, M. S. (1995). An Escherichia coli chromosomal ars operon homolog is functional in arsenic detoxification and is conserved in gram-negative bacteria. *Journal of bacteriology*, *177*(8), 2050-2056.

5. Endo, G., & Silver, S. (1995). CadC, the transcriptional regulatory protein of the cadmium resistance system of Staphylococcus aureus plasmid pI258. *Journal of bacteriology*, *177*(15), 4437-4441.

6. Park, S. J., Wireman, J., & Summers, A. O. (1992). Genetic analysis of the Tn21 mer operator-promoter. *Journal of bacteriology*, *174*(7), 2160-2171.
7. Rensing, C., Mitra, B., & Rosen, B. P. (1997). The zntA gene of Escherichia coli encodes a Zn (II)-translocating P-type ATPase. *Proceedings of the National Academy of Sciences*, *94*(26), 14326-14331.

8. Munson, G. P., Lam, D. L., Outten, F. W., & O'Halloran, T. V. (2000). Identification of a copper-responsive two-component system on the chromosome of Escherichia coli K-12. *Journal of bacteriology*, *182*(20), 5864-5871.

9. Leonhartsberger, S., Huber, A., Lottspeich, F., & Böck, A. (2001). The hydH/G genes from Escherichia coli code for a zinc and lead responsive two-component regulatory system. *Journal of molecular biology*, *307*(1), 93-105.

10. Nakano, M. M., Yang, F., Hardin, P., & Zuber, P. (1995). Nitrogen regulation of nasA and the nasB operon, which encode genes required for nitrate assimilation in Bacillus subtilis. *Journal of bacteriology*, *177*(3), 573-579.

Chapter 2. DNA copy number modulation as a tool for synthetic biology

Chapter 2.1 – Introduction to DNA modulation

A defining goal of synthetic biology is to engineer cells to coordinate tasks that often require precise temporal modulation of gene expression. While a variety of relatively small gene circuits have been constructed and characterized, their logical combination into larger networks remains a central challenge for the field. This challenge arises primarily from the lack of compatible and orthogonal elements available for predictable dynamic control of gene expression. As an alternative approach to promoter level regulation, we have characterized the use of DNA copy number regulation as an additional layer of circuit control. Specifically, we engineer colony-wide DNA cycling in the form of plasmid copy number oscillations through a modular design that can be readily adapted for use with other existing gene circuitry in single cells. We use an endonuclease from *S. cerevisae* along with quorum sensing from *A. fischeri* to produce sustained cycling of DNA plasmid concentration across a colony of *E. coli* cells. We employ the targeted endonuclease to reversibly reduce plasmid copy number and use this mechanism as the sole negative feedback component driving oscillations. Quorum sensing is used to couple the plasmid feedback system across a colony of cells, and we observe regular oscillations of GFP expression in microfluidic chambers at different colony length scales and over extended time periods. By

incorporating elements for both positive and negative copy number regulation we improve the robustness of the circuit. Finally, we use computational modeling to quantify the robust nature of the oscillations and to demonstrate that plasmid extinction is a rare event. Copy number modulation is a generalizable principle that adds a layer of control to synthetic gene circuits, enabling dynamic regulation of circuit elements without requiring specially engineered promoters.

Chapter 2.2 – DNA cycling across a bacterial population

The adaptation of DNA copy number in response to environmental pressures is a widespread mechanism found in natural systems. Copy number adjustments can be observed as relatively fixed changes, such as the evolution of ribosomal DNA tandem arrays, but have also been found to drive rapid alterations in gene expression programs[1,2]. In synthetic biology, DNA copy number has typically been used for tuning static gene expression levels. Recently, a synthetic circuit demonstrated plasmid amplification driven by changes in E. coli growth state as cultures approached saturation[3]. In this work we present strategies that allow both negative and positive plasmid copy number modulation in E. coli cells grown at a fixed density in continuous culture. We demonstrate that this allows rational design of synthetic circuits that harness plasmid copy number to dynamically control expression levels from single genes to entire gene modules.

We initially investigated an approach for reversibly repressing gene expression by lowering copy number. We found that by expressing a nuclease alongside a plasmid containing the nuclease recognition sequence, the plasmid's copy number can be temporarily reduced below its natural levels. To characterize this effect, we constructed a ColE1 origin plasmid with a constitutive promoter driving expression of a red fluorescent protein (RFP), allowing use of fluorescence measurements to estimate copy number in vivo. We also placed a unique recognition sequence for the I-SceI meganuclease on the plasmid backbone between the origin of replication and the kanamycin resistance gene. We selected I-SceI as the nuclease because it is readily expressed in many organisms and is specific for an 18-base recognition sequence, preventing off-target restriction of the genome. On a second plasmid with the compatible p15A origin, we used the native arabinose inducible promoter from E. coli to drive expression of the I-SceI protein (see Supplementary Fig. 2.1 for plasmid diagrams). In E. coli cells transformed with both plasmids, induction of I-SceI with arabinose results in a decrease of the ColE1 plasmid copy number, as evidenced by a reduction in RFP levels both in batch (Fig. 2.1, a) and in continuous culture (Fig. 2.1, b). We verified by qPCR that the copy number of the ColE1 plasmid with the recognition sequence is reduced upon induction of I-SceI (Fig. 2.1, a). Using multiple primer sets annealing to different locations around the plasmid, we found that the majority of linearized plasmids are quickly degraded (Supplementary Fig. 2.2). Expression of high levels of nuclease did not

result in any obvious effect on cell size or growth during this period, as visualized by single cell microscopy (Fig. 2.1, b: bottom). These experiments demonstrate that a targeted nuclease can be utilized to negatively regulate expression of genes on a plasmid, even those driven by unregulated promoters.

We reasoned that by controlling expression of the nuclease, an entire module of genes and promoters can be regulated when placed on a plasmid containing the cognate cut site. To demonstrate the utility of this novel mode of regulation, we constructed a synthetic gene oscillator that utilizes plasmid copy number repression by a nuclease as the negative feedback component. The architecture of the circuit (Fig. 2.1, c) is adapted from a synthetic oscillator previously constructed in our group[4], utilizing the *lux* quorum sensing system from *A. fischeri*[5]. In brief, LuxI catalyzes production of a diffusible N-acyl homoserine lactone (AHL) molecule, which binds to the constitutively produced LuxR transcription factor and activates transcription from the *luxI* promoter, thereby forming a positive feedback loop. In the synthetic oscillator circuit the *lux* quorum sensing genes *luxI* and *luxR* are placed under their native bidirectional promoter on the ColE1 origin plasmid ("activator plasmid"). An additional copy of this promoter on the same plasmid drives expression of a green fluorescent protein (GFP) and serves as a readout of *luxI* promoter activation state. A second p15A origin plasmid ("repressor plasmid") contains a third copy of the *luxI* promoter driving expression of I-SceI, which targets and represses the activator plasmid. To facilitate fast protein turnover dynamics, LuxI, GFP, and I-SceI all

have an added ssrA tag, targeting them for degradation by the native E. coli ClpXP protease[6]. When E. coli transformed with both these plasmids are grown in continuous culture in microfluidic cell chambers, we observe regular oscillations of GFP expression that are synchronized across the cells within each chamber (Fig. 2.1, d and e; see Supplementary Fig. 2.3, a for microfluidic chip design). This synthetic oscillator circuit demonstrates that nuclease-mediated copy number repression can override the strong positive feedback provided by the *lux* quorum sensing module and is a powerful tool in controlling gene expression.

Since the synthetic oscillator circuit we constructed relies on induced copy number changes in one plasmid, we measured the copy number of the second plasmid in the system as well. It has been reported that for the ColE1 family plasmids, which includes p15A, copy number can be altered by changes in cell state, including metabolic burden resulting from high rates of translation[7]. To investigate whether the copy number of p15A was affected over the course of a period of oscillation, we measured the levels of RFP expressed from a constitutive promoter on the p15A repressor plasmid. We observed a small increase in RFP signal following each peak of GFP signal (Fig. 2.2, c: top), suggesting a slight temporary amplification in p15A copy number during each period of oscillation. In response to this observation, we considered whether repressor plasmid copy number modulation could play a dynamical role in oscillations.

To explore this further, we constructed a p15A plasmid where *luxI* promoter activation directly amplifies plasmid copy number. We took advantage

of the native regulation of the p15A plasmid, which like in other ColE1 family plasmids is composed of an antisense RNA system with two convergent promoters[8]. One of these promoters is responsible for producing an RNA transcript (RNAII) that serves as a pre-primer for plasmid replication. It has been demonstrated for ColE1 plasmids that overproduction of RNAII leads to amplification of copy number[9]. We reasoned that since the p15A origin utilizes an analogous regulation mechanism[10], overproduction of RNAII on a p15A origin plasmid should similarly raise copy number. To construct such a plasmid, we removed the transcriptional terminator directly downstream of the I-SceI gene on the repressor plasmid and replaced it with a second copy of the *luxI* promoter. In this modified repressor plasmid both of the *luxI* promoters upstream and downstream of the I-SceI gene are oriented such that transcription from these promoters drives into the p15A origin in the same direction as the RNAII promoter (Fig. 2.2, a). To visualize the effect of transcription into the origin on plasmid copy number, we replaced the I-SceI gene with an ssrA-tagged *gfp* gene to serve as a reporter for *luxI* promoter activation. RFP expressed from a constitutive promoter on the same plasmid reports on copy number. As expected, we observed that *luxI* promoter activation by exogenously added AHL, seen by a rise in GFP, leads to amplification of p15A copy number and a transient increase in RFP (Fig. 2.2, b). By qPCR we measured a 2.3-fold ($p < 0.001$) amplification of p15A copy number after 90 minutes of induction with 450 nM AHL (Supplementary Fig. 2.4).

Next, we investigated the effects of this modified repressor plasmid on the oscillator dynamics. We imaged cells transformed with both the activator plasmid and the modified repressor plasmid, growing under the same conditions used in previous microfluidic experiments. Using the RFP reporter we found that the previously observed oscillations in repressor plasmid copy number are indeed amplified 1.5-fold ($p < 0.001$) in this modified oscillator circuit (Fig. 2.2, c: bottom). Introduction of RNAII overexpression also leads a decrease in period and altered waveform as compared to the first oscillator. To assess whether the use of a nuclease and RNAII overexpression adversely affects plasmid maintenance and thus the stability of this circuit, we imaged growing cells in a microfluidic device continuously for several days under constant antibiotic selection. We observed regular oscillations without apparent mutations or significant effects on cell growth (Fig. 2.2, d and e).

We compared the two oscillator circuits to investigate whether the addition of engineered copy number amplification on the repressor plasmid produces more robust oscillations. As a test for robustness with respect to microfluidic device geometries, we cultured cells of each strain in an alternate microfluidic device with cell chambers designed to have approximately 5-fold larger volume per chamber (Supplementary Fig. 2.3, b). When the chambers were seeded with the strain containing the second circuit modified to include engineered RNAII overexpression, we observed regular oscillations in these larger chambers as well (Fig. 2.3, a). Furthermore, this modified circuit produced

oscillations even when the growth medium was switched from rich lysogeny broth (LB) to minimal salts medium (M9) with glycerol (Fig. 2.3, b). In contrast, we observed at best low amplitude irregular oscillations in these larger microfluidic chambers using the original circuit with the unmodified repressor plasmid grown in either LB or M9 medium. This suggests that modifying the circuit to incorporate both negative and positive DNA copy number regulation produces more robust oscillations across different culturing conditions.

Our observations were integrated into quantitative reaction network models for the circuit without RNAII overexpression feedback (model 1) and the circuit with RNAII overexpression feedback (model 2), the two of which differ only in that model 2 contains control of the repressor plasmid copy number. These models distill our understanding of the core plasmid copy number oscillatory mechanism and are based on degrade-and-fire models previously fit to experiments[4, 11, 12], including the effects of both delay in feedback11 and proteolytic queueing[13] (see Supplementary Information for details). Parameter values for these models were determined by jointly fitting model GFP trajectories to corresponding representative mean fluorescence trajectories obtained from the oscillator experiments depicted in Figures 2.1 and 2.2. The models agree well with experimental trajectories (Fig. 2.3, c), and the models predict that oscillation amplitudes for the plasmid copy number are small enough that plasmid extinction in cells is a rare event (Fig. 2.3, d). A robustness analysis of the models supports that these oscillator designs are robust to general parameter variation

(Fig. 2.3, e, Supplementary Fig. 2.5). Curiously, model 1 was identified as being more robust than model 2. This last observation did not appear to be consistent with our experiments, but we believe the more sinusoidal oscillations in Figure 2.2, d as compared to the strong relaxation oscillations in Figure 2.1, d placed model 2 in the vicinity of a Hopf bifurcation, which would be consistent with lower robustness for model 2. Additional investigation of quasi-1D scans of robustness support model 2 as being less robust and being close to a bifurcation (Supplementary Fig. 2.6). However, we addressed the robustness experiments in Figure 2.3, a and b by demonstrating that slightly perturbed parameters consistent with a change in trap geometry could lead to a situation where only model 2 oscillates, as in the experiment (Supplementary Fig. 2.7). This last observation indicates that robustness of the circuits may be dependent on trap geometry.

The engineered interactions discussed previously are sufficient to explain oscillations, but we also considered the impact of implicit interactions arising from the limited abundance of transcription factors. When transcription factors are not present in a large excess, the ratio of binding sites to binding proteins becomes relevant to gene expression. Amplified DNA copy number necessarily implies an increased number of DNA binding sites, an effect that is multiplied when a single DNA copy contains multiple binding sites. In the oscillator circuits described in this work, positive feedback is sustained only when LuxR sufficiently binds to and activates the *luxI* promoter located upstream of the *luxI* gene. Thus

we reasoned that this positive feedback could be effectively interrupted if sufficient decoy binding sites are supplied, effectively titrating LuxR away from the promoter driving the *luxI* gene. To demonstrate this, we constructed a version of the repressor plasmid that retains both LuxR binding sites (one in each *lux* promoter), but includes a downstream transcriptional terminator to uncouple *luxI* promoter activation from p15A plasmid amplification. Instead, we used an IPTG-inducible promoter to drive amplification of this repressor plasmid. To remove the effect of I-SceI, we deleted the I-SceI cut site on the activator plasmid. With this system, we found that ongoing LuxR-mediated positive feedback from the activator plasmid can be interrupted by producing a large number of additional LuxR binding sites via amplification of the repressor plasmid. When IPTG is removed and repressor plasmid copy number is allowed to decay back to natural levels, positive feedback from the activator plasmid resumes (Supplementary Fig. 2.8). Observation of this indirect interaction due to transcription factor titration points towards future opportunities to control DNA copy number modulation in applications.

The original genetic clock[14] and toggle switch[15] circuits firmly established the engineering pillar of the field of synthetic biology. While both designs employed similar forms of transcriptional regulation, they were also prescient concerning the modern practice of parsing gene circuits into “analog” and “digital” components[16–18]. Subsequently, analog clocks evolved into platforms for exploring the synchronization of gene circuits within[4,19] and

between bacterial colonies[20,21], while digital logic was generalized to complex intracellular algorithms[22–29] and memory storage[30,31]. Our work establishes a framework for the engineering of a DNA "master clock" at the colony level that can serve to coordinate digital sub-processing within single cells. More generally, our results demonstrate how DNA copy number modulation can be used as a general tool for controlling gene expression in synthetic biology.

Chapter 2.3 – Supplementary information

Details of the Quantitative Model

We explored a number of models for the experimental oscillators. The simplest models considered the dynamics for the concentrations of LuxI, I-SceI, GFP, and the plasmid expressing LuxI, with the constraint that the concentrations of GFP and I-SceI are simply proportional to LuxI. These two-dimensional models with appropriate parameter values produced oscillations reminiscent of experimental trajectories (results not shown), which demonstrates that the basic elements of gene regulation-based positive feedback coupled to plasmid copy number regulation-based negative feedback are theoretically sufficient for oscillations.

These suggestive results prompted us to explore a more complex empirical model with additional elements, leading to a model that includes gene regulation-based positive feedback, plasmid copy number regulation-based

negative feedback, intracellular delay in feedback, and proteolytic queueing effects, all of which are known to potentially be important based on the design of the circuit and based on prior studies. These additional details led to a model that can both describe aspects of the experimental data and also be relatively robust with respect to parameter variation, as Supplementary Section 3 will discuss. It is worth noting that we found a few qualitatively similar parameter sets for the model that all fit experimental data comparably, so we picked our final parameter set based largely on robustness.

The model considers the dynamics of five key variables: the concentrations of LuxI (labeled A), I-SceI (labeled S), GFP (labeled G), the plasmid expressing LuxI (labeled PA), and the plasmid expressing I-SceI (labeled PS). Concentrations are indicated by square brackets, e.g. the concentration [A] for species A. Furthermore, we include explicit dynamics for Ai, Si, PAi, and PSi (i = 1, 2, ...5), which effectively model a delay in the production of species or perhaps a delay in the feedback on the species. These effective intermediate species were introduced, because we expect delay to exist for both protein production and plasmid production in the experimental context. Note that we use the concentration [A] as an effective proxy for the concentration (up to some scale) of other activating species, such as AHL, to simplify our system.

These species are governed by the following reactions, where the reaction velocities are assumed to have appropriate mass action terms included.

Production of precursors for activator and repressor follow from the respective reactions

$$\mathrm{PA} \xrightarrow{F([\mathrm{A}])} \mathrm{PA} + \mathrm{A}_1 \tag{1}$$

$$\mathrm{PS} \xrightarrow{F([\mathrm{A}])} \mathrm{PS} + \mathrm{S}_1 \tag{2}$$

Note that the velocity of these reactions depends linearly on plasmid concentrations due to mass action terms, which allows plasmid copy number to influence gene expression. F([A]) encodes gene regulation by the act of LuxI producing AHL, which in turn activates PLuxI promoters. F([A]) can be written

$$F([\mathrm{A}]) = \frac{\alpha \cdot \left(1 + f \cdot \left(\frac{[\mathrm{A}]+H0}{A0}\right)^n\right)}{1 + \left(\frac{[\mathrm{A}]+H0}{A0}\right)^n} \tag{3}$$

where α is the maximum production rate (per plasmid), f characterizes the strength of gene activation by A (value constrained by f ≥ 1), n is the cooperativity of gene activation (set to n = 2, but allowed to vary in our robustness analysis), A0 is the value of [A] required to strongly activate gene expression, and H0 allows for a generally time-dependent background level of AHL that can stimulate activation. We assume that the background level of H0 increases suddenly from a0 to a1 at a time t0, obeying the equation

$$H0 = a0 \cdot (1 - \theta(t - t_0)) + a1 \cdot \theta(t - t_0) \quad (4)$$

with θ the Heaviside step function.

Production of the intermediates, A1 and S1, eventually leads to the arrival of mature forms A and S, respectively, via the reactions

$$\mathrm{A_i} \xrightarrow{ktauA} \mathrm{A_{i+1}}\ , \quad i = 1,2,..,4 \quad (5)$$

$$\mathrm{A_5} \xrightarrow{ktauA} \mathrm{A} \quad (6)$$

$$\mathrm{S_i} \xrightarrow{ktauS} \mathrm{S_{i+1}}\ , \quad i = 1,2,..,4 \quad (7)$$

$$\mathrm{S_5} \xrightarrow{ktauS} \mathrm{S} \quad (8)$$

with ktauA and ktauS being rate constants that characterize the delay in production (effectively a feedback delay), with respective associated mean delay 5/ktauA and 5/ktauS. These delays appeared to be important to fit the initial large pulse of GFP seen in experiment.

The proteins A and S are tagged for rapid degradation by the protease ClpXP, so we model their degradation using enzymatic kinetics. This is modeled by the reactions

$$\mathrm{A} \xrightarrow{GA([\mathrm{A}],[\mathrm{S}])} \varnothing \tag{9}$$

$$\mathrm{S} \xrightarrow{GS([\mathrm{A}],[\mathrm{S}])} \varnothing \tag{10}$$

with functions

$$GA([\mathrm{A}],[\mathrm{S}]) = \frac{\mu \nu_1}{K + \nu_1[\mathrm{A}] + \nu_2[\mathrm{S}]} \tag{11}$$

$$GS([\mathrm{A}],[\mathrm{S}]) = \frac{\mu \nu_2}{K + \nu_1[\mathrm{A}] + \nu_2[\mathrm{S}]} \tag{12}$$

where μ is the maximum degradation velocity, and the parameters K, ν1, and ν2 characterize the affinities of protein to the protease. Recall that mass action terms should be included in the reaction velocities.

Cutting of PA is modeled by a bimolecular reaction that allows S to degrade PA

$$\mathrm{PA} + \mathrm{S} \xrightarrow{kc} \mathrm{S} \tag{13}$$

with kc the cutting rate constant. We assume then that the act of cutting immediately degrades the activator plasmid, which is likely reasonable given that

we expect linearized DNA to be degraded within the cell. It is possible that linearized DNA could re-circularize, but we do not model this. It is also possible that re-circularization may only effectively modify (reduce) the rate constant kc.

Proteins are assumed to be diluted due to cell growth and division. This is modeled by the reactions

$$\mathrm{A} \xrightarrow{ga} \varnothing \tag{14}$$

$$\mathrm{S} \xrightarrow{gs} \varnothing \tag{15}$$

with ga = gs = ln 2/30.0 min. the dilution rate. Plasmids are assumed to be degraded also by dilution, so we set

$$\mathrm{PA} \xrightarrow{k} \varnothing \tag{16}$$

$$\mathrm{PS} \xrightarrow{k} \varnothing \tag{17}$$

with k = ln 2/30.0 min. We will allow ga, gs, and k to be varied independently in our robustness analysis, even though we set them to have the same value here.

For the oscillator without the effect of RNAII overexpression on the repressor plasmid (termed Model 1), we model plasmid production by the production of an intermediate, e.g. PA1 representing a partially replicated plasmid. Production follows from the reactions

$$\varnothing \xrightarrow{k \cdot PA0} \mathrm{PA}_1 \tag{18}$$

$$\varnothing \xrightarrow{k \cdot PS0} \mathrm{PS}_1 \tag{19}$$

where the parameters PA0 and PS0 allow different plasmid copy numbers for activator and repressor plasmids. Plasmid intermediates eventually lead to an complete plasmid by the additional reactions

$$\mathrm{PA}_i \xrightarrow{ktauPA} \mathrm{PA}_{i+1}\ ,\quad i = 1,2,..,4 \tag{20}$$

$$\mathrm{PA}_5 \xrightarrow{ktauPA} \mathrm{PA} \tag{21}$$

$$\mathrm{PS}_i \xrightarrow{ktauPS} \mathrm{PS}_{i+1}\ ,\quad i = 1,2,..,4 \tag{22}$$

$$\mathrm{PS}_5 \xrightarrow{ktauPS} \mathrm{PS} \tag{23}$$

which are analogous to the corresponding reactions for A and S. Notice that we do not allow plasmid extinction using this scheme, but we will check self-

consistently that plasmid copy numbers do not become so low that extinction is likely to be an issue.

For the oscillator with the effect of RNAII overexpression affecting repressor plasmid copy number (Model 2), we replaced the production reaction for PS1 by the reaction

$$\varnothing \xrightarrow{R([\mathrm{A}])} \mathrm{PS}_1 \tag{24}$$

with

$$R([\mathrm{A}]) = k \cdot PS0 \cdot \frac{fPS1 + fPS2 \cdot \left(\frac{[\mathrm{A}]}{APS}\right)^{nPS}}{1 + \left(\frac{[\mathrm{A}]}{APS}\right)^{nPS}} \tag{25}$$

with parameters fPS1, fPS2, APS, and nPS parameters characterizing this function. We set f PS1 = 1 and nPS = 1, but we allow these parameters to vary in our robustness analysis. Note that again, we use [A] as a proxy for the concentration (to to some scale) of other activator species.

The concentration of GFP ([G]) is assumed to always be proportional to [A], though the proportionality constant is allowed to vary slightly from experiment to experiment to account for variations in lamp intensity, etc. Thus, GFP effectively does not play any important dynamical role in the model.

Finally, all concentrations are divided by a fictitious standard volume, V0. We normalize this volume to V0 = 1, though this parameter is allowed to vary in our robustness analysis.

Fitting the Model Using COPASI

We encoded our model into the simulation package COPASI for basic simulation and fitting. COPASI provides several advantages: an environment that is readily installed on a variety of platforms, an ability to import from and export to a variety of model formats, and inclusion of numerous tasks (including model fitting) that can be run on models. We have shared both a COPASI model file and the corresponding SBML file as supplementary files.

The numerous unknown parameters in the model prompted us to attempt automated fitting of the time-course data to determine these parameters. Within COPASI, we fitted the model to experimental data by (1) importing two representative experimental mean GFP trajectories (averaging across a whole microfluidic trap) for the circuits without and with RNAII overexpression feedback, (2) configuring COPASI's Parameter Estimation task to simultaneously fit these two experimental results to our model, (3) ensuring the model has run to near steady state before experimental data is compared to the model (by allowing the first pulse of the experimental oscillator to be occur at a late time, roughly at 300 min.), and (4) running the Parameter Estimation task with given constraints on parameter sizes and with a given fitting method. We used

deterministic integration to simulate the model when comparing it to experimental data, using the LSODA method with relative tolerance 10^{-6} and absolute tolerance 10^{-12}. Fitted parameters followed by attempting to minimize the root-mean-square error between the model's trajectory for GFP concentration and the measured mean GFP in experiment. Recall that we assume GFP concentration is proportional to [A].

The above step (4) was executed many times to find a model that optimized accuracy. First, global optimization techniques, such as the COPASI method "Evolutionary Strategy," were used to find broad regions of parameter space with accurate solutions. These methods were followed with local optimization methods in COPASI, such as "Hooke and Jeeves," to further increase accuracy. We ultimately found a deterministic ODE model that was both accurate and exhibited reasonable robustness, the latter of which is discussed in the next section. We use the term "Model 1" for the model where RNAII overexpression does not affect repressor plasmid copy number, while we use the term "Model 2" for the model where RNAII overexpression does affect repressor plasmid copy number. These models share the same parameters, except that Model 2 has additional parameters to characterize the potential effects of RNAII overexpression.

Our final parameter set for our model fit is reported in Supplementary Table 2.2.

Robustness Analysis of the Quantitative Model

We checked whether our model was robust, i.e. it produces oscillations for a wide range of parameter values. For models with few (~ 2) parameters, robustness can be checked using bifurcation diagrams. However, our model has many parameters, and a standard bifurcation analysis may not properly reflect any underlying robustness. To address these concerns, we performed a robustness analysis that samples many parameter sets around our fitted parameter set, and we tested whether these parameter sets corresponded to oscillations. Custom Python software was written to translate C-code output from COPASI into fast Python code, since COPASI was apparently unable to perform the robustness analysis we desired. The scipy.integrate.odeint module in the scipy library was used, with relative tolerance 10−6 and absolute tolerance 10−12.

Robustness was explored as follows. We scanned a parameter η from 0.00 to 0.80 in increments of 0.02. For each value of η, we constructed 100 parameter ensembles that each consisted of 1000 random parameter sets. A random parameter set is generated by multiplying the respective fitted value of every parameter in Supplementary Table 2 by an independent random number uniformly distributed between 1 − (η /2) and 1 + (η /2). This samples a large hypervolume of parameter space that is difficult to obtain using bifurcation diagrams. A trajectory was determined to be oscillatory if the standard deviation

(across time) of the trajectory [A](t) exceeded 10% of the mean (across time) during a specific window of time. For simulations starting at time t = 0, this window of time was 814-1512 min. for Model 1 (without RNAII overexpression effects) and 862-1510 min. for Model 2 (with RNAII overexpression effects). The time each model was induced was t0 = 288 min.

Figure 2.3, e reports the robustness based on this analysis. These statistics support that Model 1 is generally more robust than Model 2 with respect to oscillations. We suspect this is due to Model 2 being tuned closer to a bifurcation such as a Hopf bifurcation, which would lead to GFP oscillations that appear more sinusoidal in shape, i.e. not relaxing to zero during each oscillation. The trajectories of mean GFP for the experimental construct with RNAII overexpression regulation exhibit similar oscillations as in Model 2, i.e. oscillations that do not relax to zero (background) intensity. Thus, our numerical results and this overall logic suggest that the experimental construct with RNAII overexpression regulation should be less robust.

We quantified the most sensitive parameters with respect to robustness using the same data set as above (the scan of η). For each ensemble of 1000 trajectories, we construct a non-oscillatory set consisting only of parameters not leading to oscillations (by our above test). Within this non-oscillatory set, the principal component with the lowest coefficient of variation (standard deviation divided by the mean) was determined. The idea is that the most sensitive parameters would have a narrow distribution in the non-oscillatory set, since a

parameter that does not heavily influence the stability of oscillations would have a wide distribution (sampling most of the default fractional range $1-(\eta/2)$ and $1+(\eta/2)$). We report out results from this principal component analysis in Supplementary Figure 2.5. We found that the degradation velocity μ was a key sensitive parameter in this investigation, which is not terribly surprising given the important role of degradation in oscillatory dynamics.

Chapter 2.4 – Figures and tables

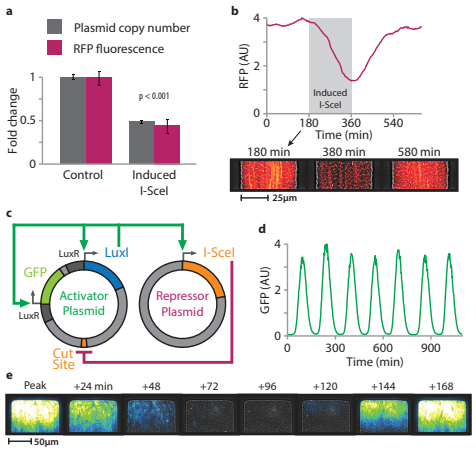


Figure 2.1. A meganuclease (I-SceI) serves as a negative feedback element in this synthetic quorum oscillator by targeting the activator plasmid. (a) To demonstrate copy number modulation I-SceI is placed under control of an arabinose-inducible promoter, with I-SceI targeting a ColE1 origin plasmid that codes for constitutive RFP production. After 3 hours of exponential growth in a flask containing arabinose, both RFP levels and ColE1 copy number have dropped significantly ($p < 0.001$ for each). Plasmid copy number was quantified by qPCR with primers spanning the cut site. Mean and SEM are displayed and p-values as calculated by independent 2-sample t-test with $n = 10$ replicates for each condition. See Supplementary Figure 2.1, a for plasmid diagrams. (b) Strain from panel a shown growing in a microfluidic cell chamber. RFP expressed from a constitutive promoter on the ColE1 origin plasmid drops sharply upon induction of I-SceI with arabinose (shaded region), then recovers upon removal. Images below show composite of phase contrast and RFP fluorescence at indicated time points. (c) Diagram of the two-plasmid circuit: a ColE1 origin plasmid (left, "activator plasmid") includes the *lux* quorum sensing system from *A. fischeri*, which serves as a positive feedback loop for synchronized *luxI* promoter activation. A p15A origin plasmid (right, "repressor plasmid") has the same quorum activated *luxI* promoter controlling expression of I-SceI, which targets the activator plasmid and thereby lowers copy number. LuxI, GFP, and I-SceI proteins are ssrA-tagged for ClpXP degradation. (d) Time series of average GFP signal from a representative chamber shows regular synchronized oscillations produced by the circuit. (e) Film strip showing composite of phase contrast and GFP fluorescence produced by the oscillator circuit in a single microfluidic chamber, covering approximately one period of oscillation.

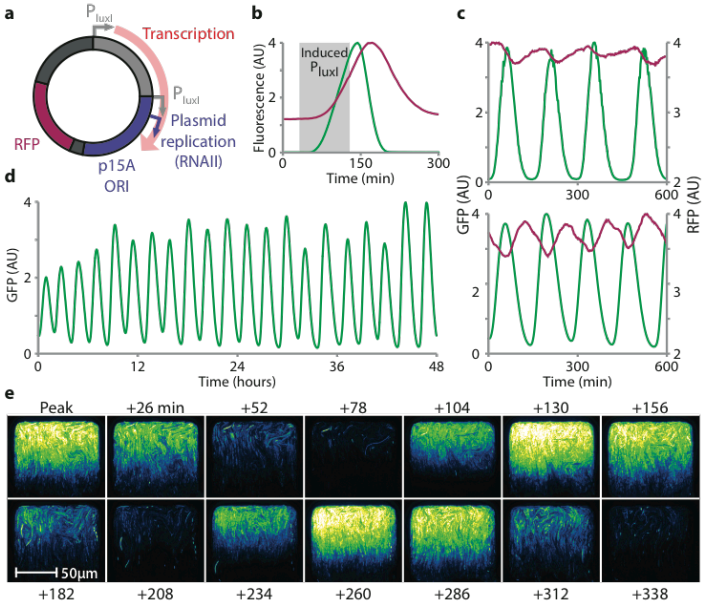


Figure 2.2. Amplification of DNA copy number by engineered transcription into the origin of replication. (a) Diagram depicting the arrangement of promoters that allows overproduction of RNAII. A second *luxI* promoter replaces the terminator after the I-SceI gene, with both promoters facing the same direction as the native p15A promoter that primes plasmid replication (RNAII promoter). Constitutive RFP production reports on plasmid copy number. (b) To demonstrate the effect of *luxI* promoter activation on copy number, the I-SceI gene is replaced with a *gfp* gene. Induction with 450 nM AHL for 90 minutes (shaded region) results in GFP production from the *luxI* promoter (shown in green) and a concomitant rise in RFP signal (shown in red), indicating plasmid copy number amplification. Note that GFP is ssrA tagged for degradation while RFP is untagged. (c) Small oscillations in RFP constitutively expressed from the p15A plasmid (shown in red) are apparent even in in the circuit without RNAII overproduction (top, circuit as shown in Fig. 2.1, c). When the repressor plasmid is modified to include RNAII overproduction, these RFP oscillations are magnified 1.5-fold (bottom). In both circuits GFP (shown in green) is expressed from the *lux* promoter on the activator plasmid. (d) Time series of average GFP signal in a representative cell chamber shows regular synchronized oscillations produced by the circuit with the RNAII overproduction. (e) Film strip of GFP fluorescence in a single chamber covering approximately 2.5 periods of oscillation, produced by the same circuit as in panel d.

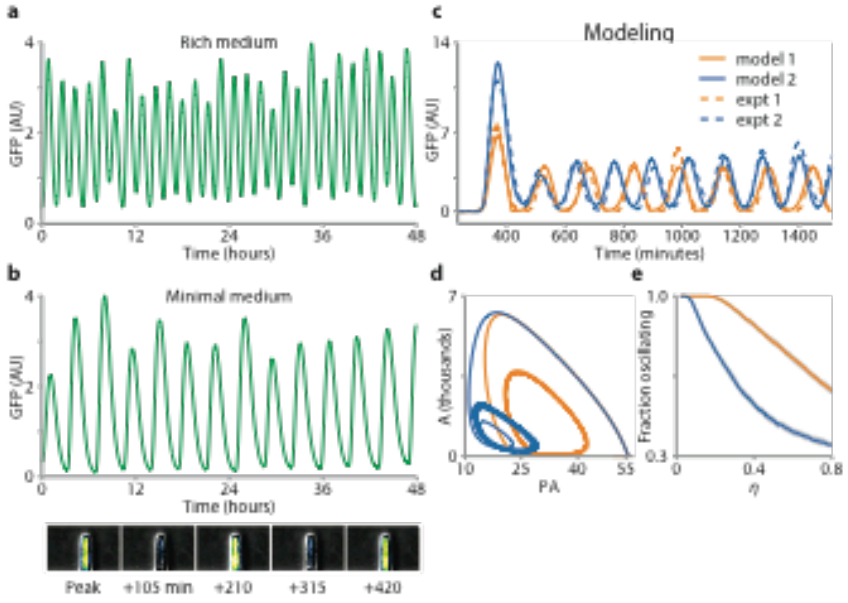
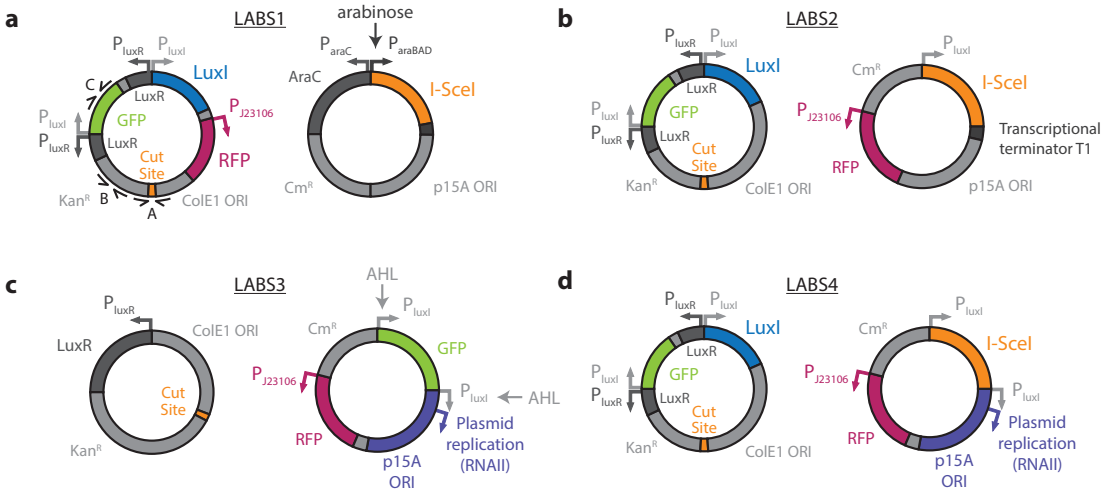
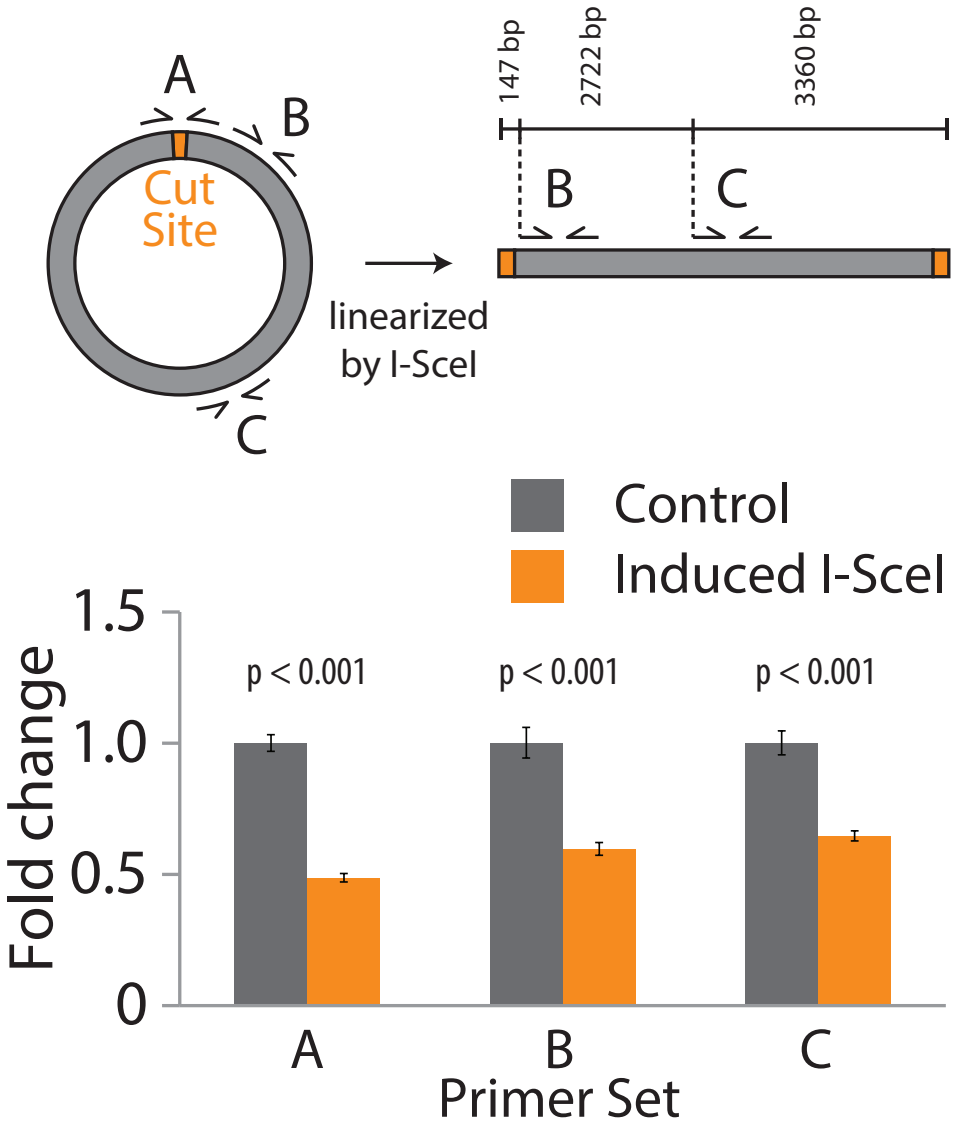


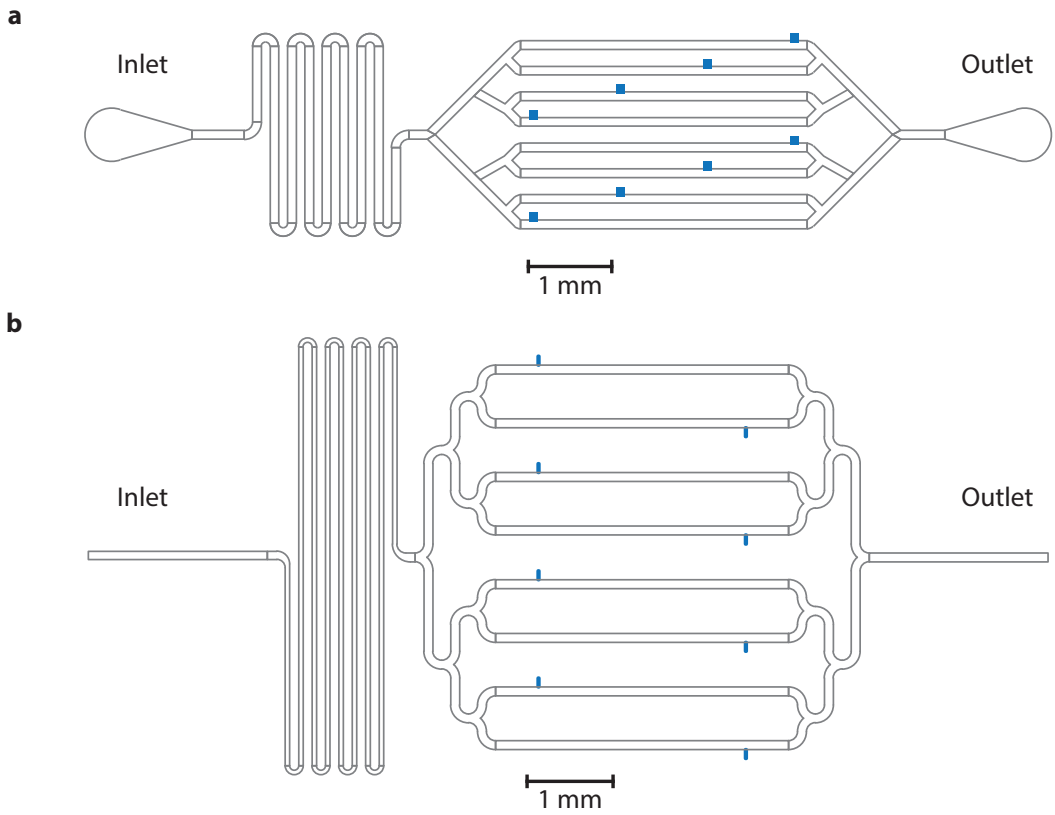
Figure 2.3. Robustness and model analysis of the DNA feedback circuit. (a) Cells containing the circuit modified for repressor plasmid amplification produce regular oscillations in an alternate microfluidic device design that has taller chambers with larger volume. (b) Same microfluidic de
vice and bacterial strain from panel a cultured with minimal salts medium instead of rich medium. Oscillations remain regular but shift to a longer period. Filmstrip below shows a composite of phase contrast and GFP fluorescence in a single chamber over two periods of oscillation. (c) Trajectories that show good agreement between mean GFP intensity from our two models (solid lines, with model 1 lacking RNAII overexpression feedback in orange, and model 2 with RNAII overexpression feedback in blue) to corresponding mean GFP intensities averaged across single representative microfluidic traps (dashed lines). Experiments have been aligned in time to have a similar time for their first peak. (d) Projection of these oscillations onto LuxI concentration (A) and plasmid copy number for activator (PA). The model predicts that plasmid copy number does not require excessive variation, which is important to avoid plasmid extinction. (e) Robustness analysis of our model fit as a function of parameter perturbation strength η suggests our model is reasonably robust in a high dimensional parameter space (see Supplementary Information). Solid lines show the fraction of perturbed parameter sets that oscillate out of 100,000 parameter sets per η value. Statistical error is comparable to the width of the line, as determined by bootstrapping over 100 independent batches of 1000 parameter sets.



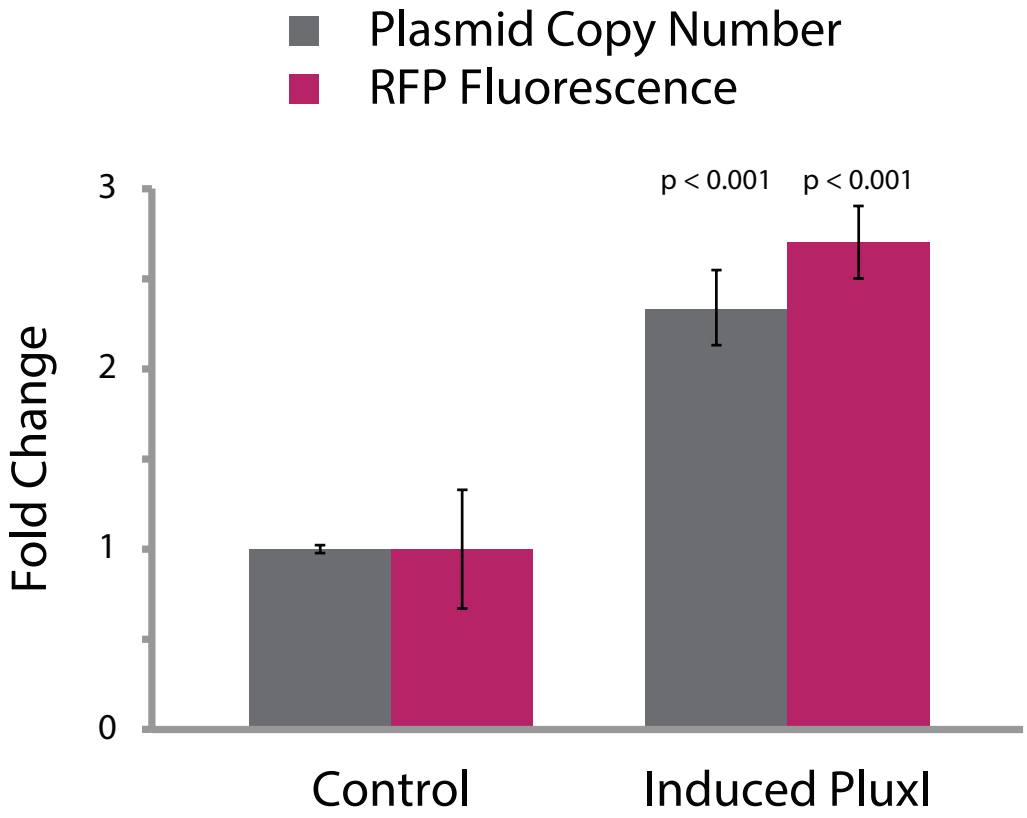
Supplementary Figure 2.1. Plasmid diagrams of primary strains used in this work. Refer to Supplementary Table 1 for additional information and full list of strains. Diagrams are not to scale. (a) Strain used to demonstrate and quantify ColE1 copy number repression by I-SceI. A, B, and C refers to primer sets used for qPCR as shown in Supplementary Figure 2.2. (b) Plasmid copy number oscillator strain without RNAII overexpression. Note the transcriptional terminator downstream of I-SceI preventing transcription from progressing into the p15A origin. (c) Strain used to visualize and quantify p15A copy number amplification by RNAII overproduction. The transcriptional terminator before the p15A origin has been replaced with a second copy of the *luxI* promoter, and the I-SceI gene with a *gfp* gene. (d) Plasmid copy number oscillator strain modified to include RNAII overproduction under control of the *luxI* promoter as in LABS3.



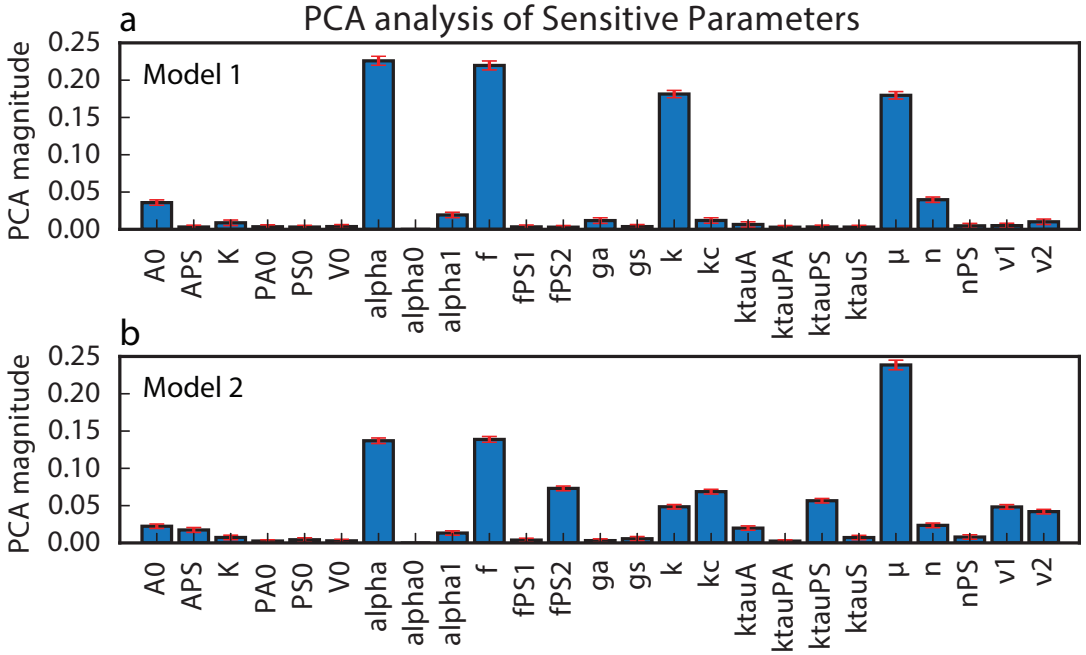
Supplementary Figure 2.2. qPCR analysis of ColE1 copy number repression using multiple primer sets annealing at different locations around the plasmid demonstrates that after being cut by I-SceI, the majority of linear plasmids are quickly degraded. Primer set A is the same that used to quantify copy number in Figure 2.1, a. See Methods for detailed description of experimental design and Supplementary Figure 2.1 for detailed plasmid diagrams.



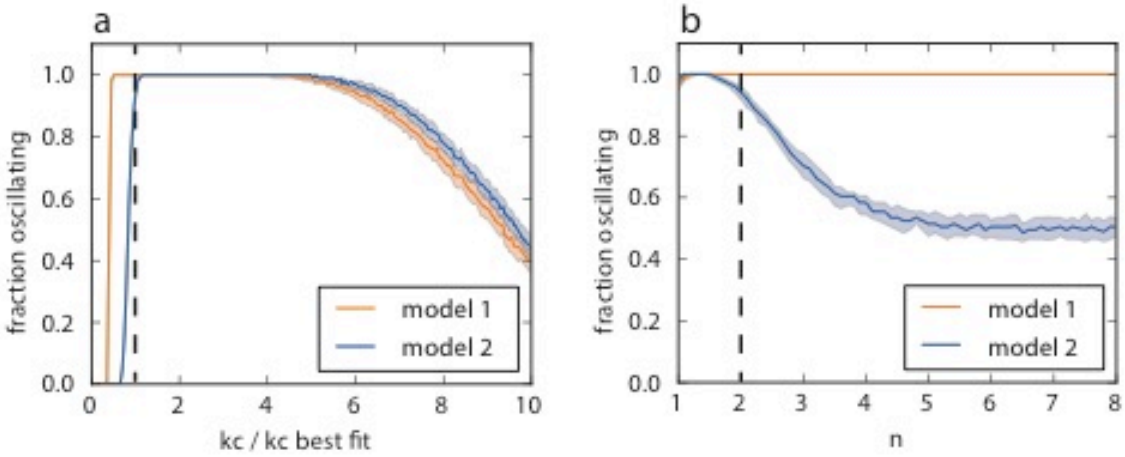
Supplementary Figure 2.3. Schematics of microfluidic devices. Each flow channel feeds growth medium to a single cell chamber (displayed in blue), preventing AHL diffusion between individual chambers. (a) Flat rectangular chambers (x, y, z = 100 μm, 85 μm, 1.6 μm) allow visualization of single *E. coli* cells. Flow channels are 30 μm high. (b) Larger and taller chambers (x, y, z = 15 μm, 100 μm, 50 μm) allow exploration of circuit dynamics in an alternate geometry. Flow channels are the same height as chambers at 50 μm.



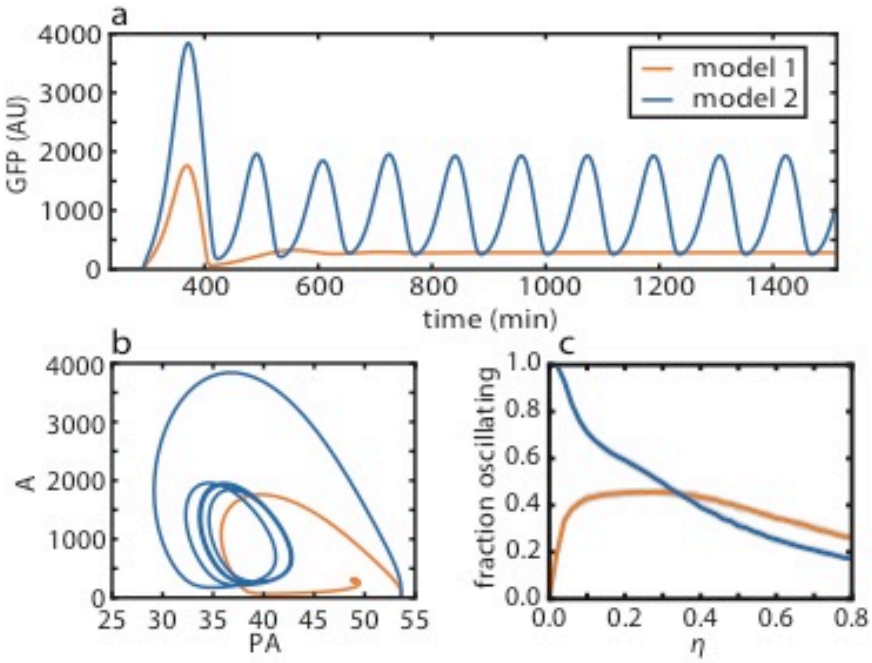
Supplementary Figure 2.4. qPCR analysis of p15A copy number amplification by RNAII overexpression. Mean values and SEM are displayed with p-values as calculated by independent 2-sample t-test with n = 10 replicates for each condition. See Methods for detailed description of experimental design and Supplementary Figure 2.1 for detailed plasmid diagrams.



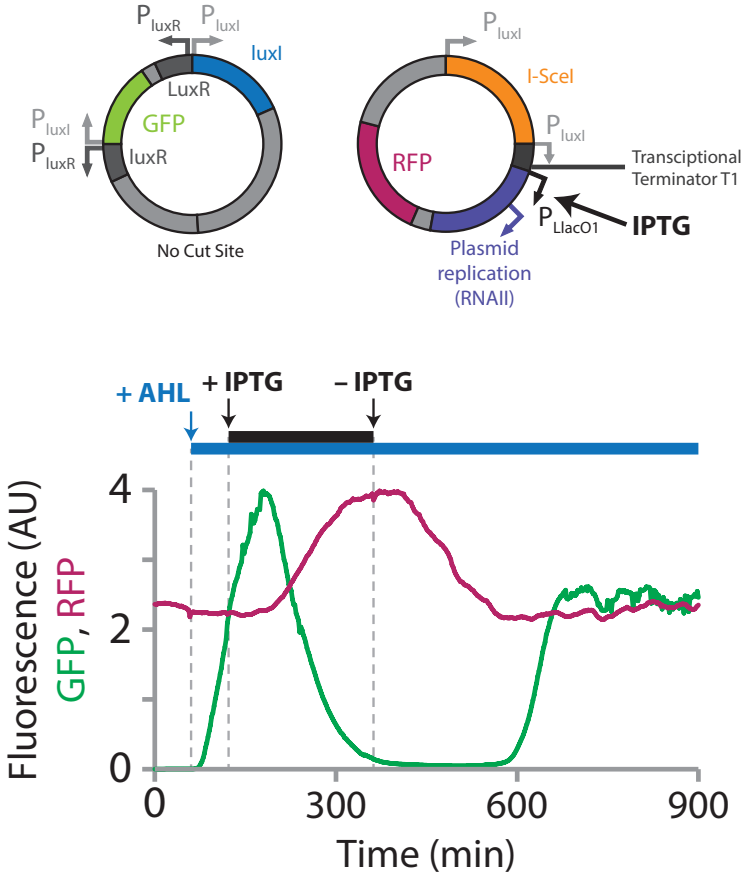
Supplementary Figure 2.5. Robustness analysis was done to identify which parameters most sensitively control oscillations (see SI for details). For (a) model 1 and (b) model 2, we formed non-oscillatory parameter sets from the data used to generate Figure 2.3, e, and we determined the principal components with the lowest coefficient of variation (a proxy for sensitivity in this set). We examined representative non-oscillatory sets using $\eta = 0.3$ and $\eta = 0.1$ when studying model 1 and model 2, respectively, containing non-oscillatory parameter set sizes of 5869 and 6468, respectively. The principal component with the least coefficient of variation was found using standard techniques, the absolute magnitude for each parameter value was taken, and the sum of principal components was normalized to 1. For each model, this process was repeated using 1000 ensembles containing a random 20% of the full non-oscillatory set. We applied bootstrapping to determine the mean and standard deviation of the principal component with the least coefficient of variation (blue bars represent the mean, red lines represent standard deviation). Overall, the enzymatic velocity μ was consistently a sensitive parameter.



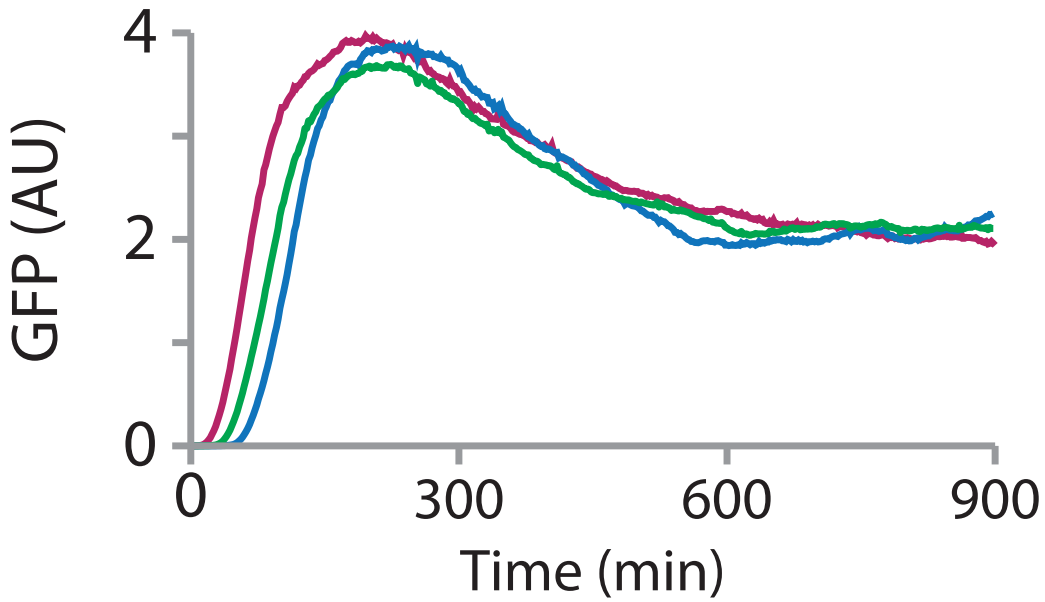
Supplementary Figure 2.6. Fraction of oscillatory models for a scan of two oscillator parameters. (a) The cutting rate constant, kc, for I-SceI was varied from 0 to 10-times its best fit value. All other parameters were varied using a uniform distribution ranging ±5% (similarly as done for Figure 2.3, e for η = 0.10, but using 100 ensembles of size 100 each to estimate error). The dashed line indicates the best fit value. (b) We similarly investigated robustness for the cooperativity coefficient n, which was scanned from 1 to 8. The observation that model 2 for both (a) and (b) has fewer nearby parameter sets that oscillate in the vicinity of the best fit is consistent with the picture that model 2 is tuned closer to a bifurcation point.



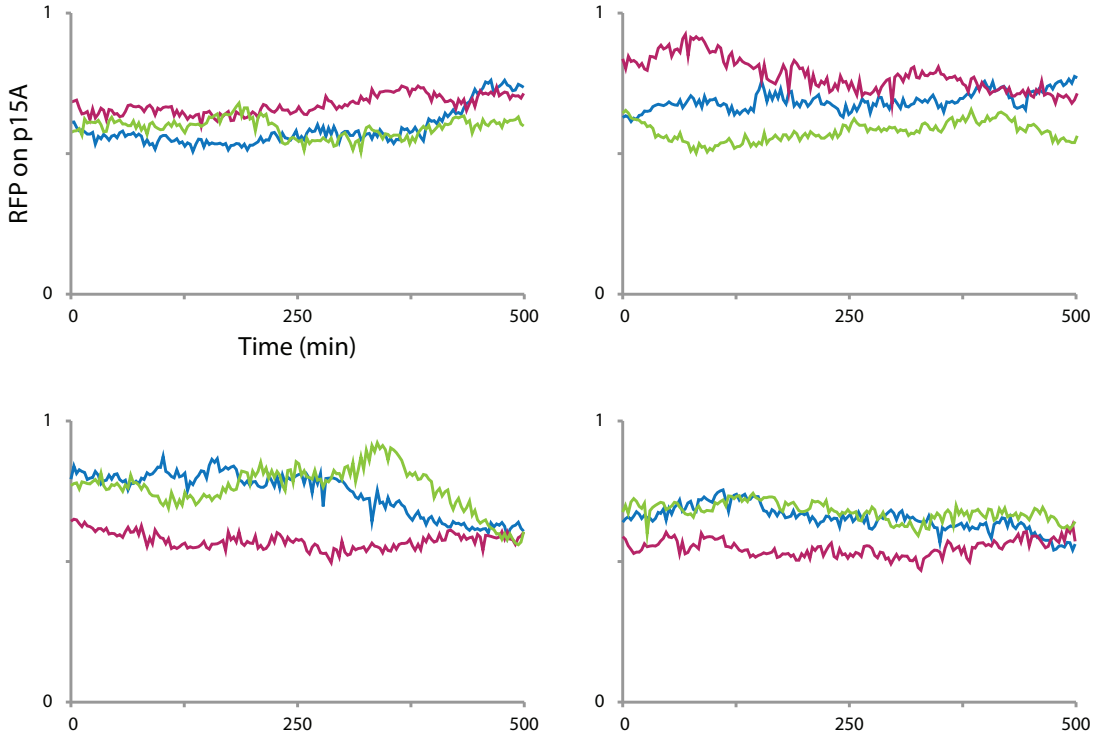
Supplementary Figure 2.7. To address the observations in Figure 2.3, where only the circuit with RNAII overexpression feedback oscillated, we examined our model for parameters modified to reflect a change in trap geometry, e.g. leading to slower AHL buildup in the trap. In particular, we effectively modified the level of LuxI needed to activate the circuit by increasing the parameter value for A0 by 2-fold, and we modified the delay for positive feedback by increasing the parameter value for ktauA by 6-fold. (a) Model 1 did not exhibit oscillations after a short transient, while model 2 exhibited sustained oscillations, as is consistent with experiment. (b) Plasmid copy numbers maintained reasonable values, suggesting plasmid extinction should not be a concern. (c) These solutions were reasonably robust with respect to general parameter variation. Robustness analysis was performed as in Figure 2.3, e.



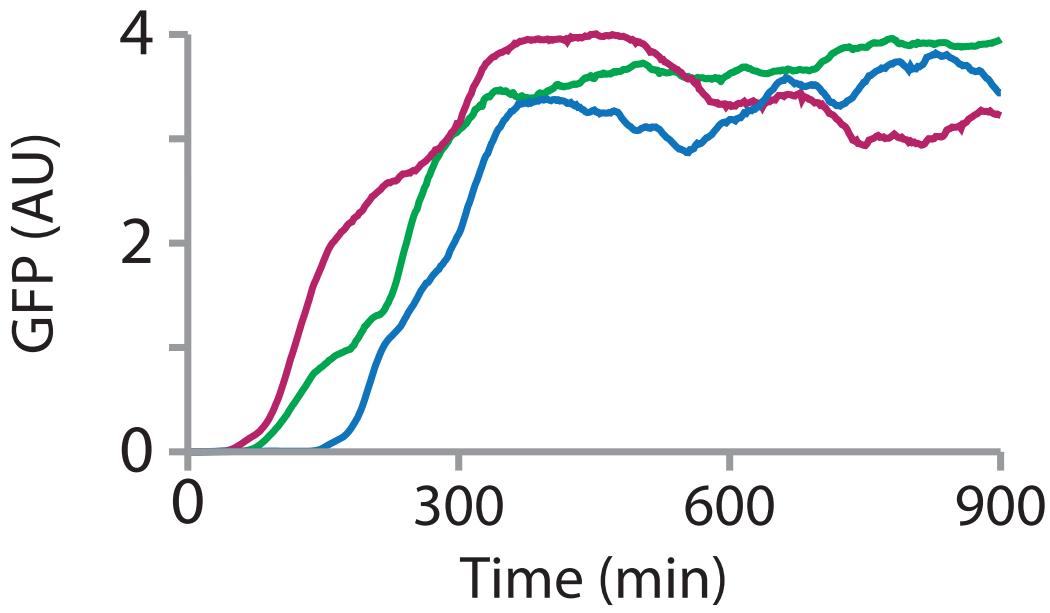
Supplementary Figure 2.8. p15A copy number amplification counteracts positive feedback from the activator plasmid even without cutting by I-SceI. The activator plasmid used here lacks an I-SceI cut site. The repressor plasmid has been modified to allow inducible plasmid amplification with IPTG, driven by the lac-repressible PLlacO1 promoter. Both *luxI* promoters are left intact, each containing a single LuxR binding site, however the transcriptional terminator downstream of the *luxI* promoters prevents transcription from progressing into the p15A origin. RFP reports on p15A copy number and GFP on *lux* activation state. A small amount of AHL (5 nM) is introduced at t = 1 hour to start *lux* positive feedback from the activator plasmid, causing a rise in GFP signal. At t = 2 hours, 100 μM IPTG is added in addition, causing amplification of the p15A copy number as seen by the rising RFP signal. IPTG is removed again at t = 6 hours, which allows p15A copy number to slowly drop back to natural levels by dilution due to cell division.



Supplementary Figure 2.9. Replacing the I-SceI gene on the repressor plasmid in the oscillator circuit with an inactive variant of I-SceI abolishes oscillations. Average GFP signal across three individual representative chambers is plotted, demonstrating that nuclease activity by a functional I-SceI protein is required to produce oscillations. Data shown is from strain LABS6, which is the same as LABS2 with the only modification to the circuit being a single amino acid change (D44A) in the coding sequence of I-SceI. Cells were grown in the same conditions as the functioning oscillator strains.



Supplementary Figure 2.10. Oscillations in RFP signal from the repressor plasmid do not originate from native regulation of p15A plasmid copy number alone. Data shown is from strain LABS7, transformed with the p15A repressor plasmid and lacking the activator plasmid. RFP is driven by a constitutive promoter. Each trajectory represents a manually tracked single cell growing in rich medium.



Supplementary Figure 2.11. The activator plasmid alone produces a sustained high level of GFP. Data shown is from strain LABS8 transformed with the activator plasmid alone and lacking a p15A plasmid. Cells were grown in the same conditions and microfluidic device as those used to produce oscillations with other strains. Each trajectory represents the average GFP signal across an individual microfluidic chamber.

Supplementary Table 2.1. Table of strains used in this work and gene expression constructs found on each plasmid. All experiments were conducted in *E. coli* MG1655. LABS8 is in strain MG1655-z1 which is modified to have high levels of constitutive lac repressor expression from the genome. RFP refers to mKate2 and GFP refers to sfGFP. Constructs with promoter "Plux" and denoted "+ *luxR*" contain the entire bi-directional *lux* promoter (both PluxI and PluxR) and the *luxR* gene. J23106 is a medium-strength constitutive promoter. PLlacO1 is an IPTG-inducuble (i.e. LacI-repressible) promoter. SsrA tags used here code for the amino acids AANDENYALAA and are inserted directly before the stop codon. I-SceI D44A refers to an inactive variant of I-SceI which lacks nuclease activity. Plasmid diagrams of LABS1-4 are available in Supplementary Figure 2.1. A diagram of LABS5 is available in Supplementary Figure 2.8. LABS6-8 are variants used for control experiments and are similar to previous strains as noted.

Strain	Construct(s) on ColE1 origin plasmid	Construct(s) on p15A origin plasmid	Referenced in Figure
LABS1	Plux-*luxI*-ssrA + *luxR* Plux-*gfp*-ssrA + *luxR* pJ23106-*rfp* (no ssrA)	ParaBAD-*ISceI* (no ssrA)	Fig. 1, a and b Extended Data Fig. 2
LABS2	Plux-*luxI*-ssrA + *luxR* Plux-*gfp*-ssrA + *luxR*	PluxI-*ISceI*-ssrA PJ23106-*rfp* (no ssrA)	Fig. 1, c and d Fig. 2, c: top
LABS3	PluxR-*luxR*	PluxI-*gfp*-ssrA PJ23106-*rfp* (no ssrA) (with copy amplification: PluxI-RNAII)	Fig. 2, a and b Extended Data Fig. 4
LABS4	Plux-*luxI*-ssrA + *luxR* Plux-*gfp*-ssrA + *luxR*	PluxI-*ISceI*-ssrA PJ23106-*rfp* (no ssrA) (with copy amplification: PluxI-RNAII)	Fig. 2, c: bottom, d, and e Fig. 3, a and b
LABS5	Plux-*luxI*-ssrA + *luxR* Plux-*gfp*-ssrA + *luxR* (as in LABS2,4)	PluxI-$ISceI_{D44A}$-ssrA PJ23106-*rfp* (no ssrA) (as in LABS2, but with nonfunctional I-SceI)	Extended Data Fig. 7
LABS6	n.a.	PluxI-*ISceI*-ssrA PJ23106-*rfp* (no ssrA) (with copy amplification: PluxI-RNAII) (as in LABS4)	Extended Data Fig. 8
LABS7	Plux-*luxI*-ssrA + *luxR* Plux-*gfp*-ssrA + *luxR* (as in LABS2,4)	n.a.	Extended Data Fig. 9
LABS8	Plux-*luxI*-ssrA + *luxR* Plux-*gfp*-ssrA + *luxR* (no I-SceI cut site)	PluxI-*ISceI*-ssrA + 2nd PluxI (terminated) PJ23106-*rfp* (no ssrA) (with IPTG-inducible copy amplification: PLlacO1-RNAII)	Extended Data Fig. 10

Supplementary Table 2.2. Table of parameter values for the model fit used in Figure 2.3, c-e.

Name	Description	Value	Name	Description	Value
A0	parameter for the production rate of LuxI and I-SceI	1105.9400	gs	dilution rate of I-SceI	0.0231
APS	parameter for RNAII overexpression copy number control	929.4370	k	dilution rate of plasmid	0.0231
K	parameter for enzymatic degradation	100.7700	kc	I-SceI cutting rate constant	4.8375×10^{-5}
PA0	scale of activator plasmid copy number	53.7116	ktauA	effective delay rate constant for LuxI production	0.1810
PS0	scale of repressor plasmid copy number	11.3784	ktauPA	effective delay rate constant for activator plasmid production	0.2141
V0	volume of cell in natural units	1.0	ktauPS	effective delay rate constant for repressor plasmid production	0.1155
a0	the background level of AHL before induction (AU)	0.0	ktauS	effective delay rate constant for I-SceI production	0.3536
a1	the background level of AHL after induction (AU)	2296.6800	μ	enzymatic degradation velocity of proteins	698.8910
α	maximum production rate per plasmid for LuxI and I-SceI	0.6582	n	cooperativity parameter for LuxI and I-SceI production	2
f	fold activation for saturating AHL vs. absent AHL in for LuxI and I-SceI production	27.7632	nPS	cooperativity parameter for RNAII overexpression copy number control	1
fPS1	parameter for RNAII feedback	1.0	v_1	parameter for enzymatic degradation	3.4699
fPS2	parameter for RNAII feedback	3.4870	v_2	parameter for enzymatic degradation	4.9963
ga	dilution rate of LuxI	0.0231			

Chapter 2.5 – Protocols and methods

Strains and Culturing

All plasmids were constructed by Gibson assembly followed by transformation into Mach1 (Invitrogen) chemically competent E. coli. Plasmids were verified by Sanger sequencing prior to transformation into E. coli strain MG1655. For strains containing quorum sensing constructs, growth on plates was limited to a maximum of 10 hours and 0.2% w/v glucose was added to all plates and media used during cloning and transformation to prevent auto-induction. All experiments were conducted in the MG1655 strain using the appropriate antibiotics: 50 μg/mL kanamycin for strains containing ColE1 origin plasmids, and 34 μg/mL chloramphenicol for strains containing p15A origin plasmids. For the experiments with strain LABS1 involving induction of I-SceI with arabinose, cells were cultured in low-density conditions that prevent auto-induction. The AHL inducer used in experiments was 3-oxo-C6-HSL. A detailed description of all strains is available in Supplementary Table 2.1 and the corresponding plasmid maps are shown in Supplementary Figure 2.1. The doubling time of oscillator strains LABS2 and LABS4 was measured in a 96-well plate in rich LB medium with antibiotics as 20.7 minutes (n = 6 wells for each strain, SEM = 0.2 min).

Microfluidics and Microscopy

For each experiment, the appropriate E. coli strain was seeded from a -80◦C glycerol stock into 3 mL of lysogeny broth (LB) supplemented with 0.2% w/v glucose, the appropriate antibiotics, and 0.075% w/v Tween 20. After growth for 8-12 hours at 37◦C in a shaking incubator, the culture was diluted 100-fold into 3 mL of the same medium and grown for one additional hour. This culture was concentrated by centrifugation at 5000 rcf for 1 minute and resuspended in 3 µl of the same medium. Cells were loaded using degas-driven flow into microfluidic chambers via the waste port. Microfluidic experiments were conducted using media supplemented with appropriate antibiotics and 0.075% w/v Tween 20. Experiments involving *lux*-mediated positive feedback were done with media containing a low background concentration of AHL, which is also required to induce oscillations: 5 nM AHL for experiments using rich LB medium, and 1 nM AHL for experiments using minimal medium. The minimal medium used was M9 salts medium supplemented with 0.4% v/v glycerol, 0.2% w/v casamino acids, and 1 µM thiamine. For experiments involving induction with chemical inducers, the medium source was switched manually at the device inlet at the times indicated. Time lapse images were acquired on a Nikon TI microscope fitted with a Lumencor SOLA SE light engine for fluorescence imaging. Average fluorescence values across each chamber were determined using Fiji software[32]. A single baseline fluorescence value was measured from a region outside the chamber for each time series and used for background

subtraction. Schematics of microfluidic devices are available in Supplementary Figure 2.3.

For comparison of oscillation amplitude of RFP in the two circuits, RFP time series were detrended by subtracting a moving average over a window equal to the period of each circuit (152 min for the first circuit, 128 min for the modified circuit). De-trended time series were used to measure peak-to-trough amplitudes for each time series, which were averaged and normalized by the mean RFP fluorescence for the time series. Statistical significance was determined by independent 2-sample t-test using mean-adjusted average amplitude values for each time series with n = 40 peaks from 7 time series for the first circuit and n = 42 peaks from 6 time series for the modified circuit.

Relative Copy Number Determination by qPCR

A culture was seeded from a -80◦C glycerol stock into 3 mL LB medium supplemented with 0.2% w/v glucose. After 12 hours of growth at 37◦C in a shaking incubator, the culture was diluted 100-fold into fresh LB without glucose and grown for an additional 2 hours to ensure that cells were in the exponential growth phase. The culture was then diluted into 250 mL flasks with 50 mL pre-warmed LB, and chemical inducer was added: 0.2% w/v arabinose for copy number repression with strain LABS1 (sterile water for control), and 450 nM AHL for copy amplification with strain LABS3 (DMSO for control). After growth in the presence of inducer (3 hours for LABS1, 1.5 hours for LABS3), 100 µl samples

were taken from each flask and heated at 95◦C for 10 minutes followed by immediate freezing at -20◦C for subsequent qPCR analysis, as described in Škulj et al.[33] For fluorescence measurements, cultures were normalized to the same OD600, pipetted into 10 replicate wells in a plate, and readings were taken using a Tecan Infinite M200Pro. A culture of wild-type MG1655 at the same OD600 was used to measure background signal. Cultures were in the exponential growth phase at the time of sampling, with OD600 readings of 0.1-0.2.

qPCR primers were validated using a 5-point template dilution series to ensure that amplification efficiency for each pair was >90%. The genomic reference primers chosen were as described in Škulj et al.[33] (GCGAGCGATCCAGAAGATCT / GGGTAAAGGATGCCACAGACA). Three sets of primers were designed targeting the ColE1 plasmid in different locations: primer set A with one primer on either side of the I-SceI cut site (GACGCTCAGTGGAACGAAA / GTAATGACCTCAGAACTCCATCTG), primer set B beginning 147 base pairs from the cut site in the kanamycin resistance gene (CTCGTCAAGAAGGCGATAGAAG / CGTTGGCTACCCGTGATATT), and primer set C beginning 2722 base pairs from the cut site in the *gfp* gene (CCATTACCTGTCGACACAATCT / GTGTAATCCCAGCAGCAGTTA). The same genomic reference primer set and primer set C were also used in a separate experiment to measure p15A plasmid amplification. For relative copy number determination, cell culture samples were thawed, diluted 2000-fold in sterile water, and added at a ratio of 2:3 to a master mix prepared by adding

appropriate primers (500 nM final concentration) to Bio-Rad iQ SYBR Green Supermix. Each reaction condition was pipetted into 10 replicate wells and threshold values were determined using a Bio-Rad CFX Connect Real-Time PCR Detection System with the associated software. Melt curves were examined for each reaction well to verify the presence of a single amplification product. Relative copy numbers were calculated using the ΔΔCt method and error was calculated by propagation of the standard error of the Ct values.

Quantitative Modeling

Model creation and fitting was done using the package COPASI[34] and custom python scripts. Details concerning the quantitative modeling appear in the Supplementary Information.

Chapter 2.6 – Acknowledgements

Chapter 2, in full, has been submitted for publication as it may appear in Nature Genetics, 2017, Baumgart, Leo; Mather, Will; Hasty, Jeff. The dissertation author was the primary author of this paper.

Chapter 2.7 – Works cited in chapter 2

1. Slager, J., Kjos, M., Attaiech, L., & Veening, J. W. (2014). Antibiotic-induced replication stress triggers bacterial competence by increasing gene dosage near the origin. *Cell*, *157*(2), 395-406.

2. Narula, J., Kuchina, A., Dong-yeon, D. L., Fujita, M., Süel, G. M., & Igoshin, O. A. (2015). Chromosomal arrangement of phosphorelay genes couples sporulation and DNA replication. *Cell*, *162*(2), 328-337.

3. Marguet, P., Tanouchi, Y., Spitz, E., Smith, C., & You, L. (2010). Oscillations by minimal bacterial suicide circuits reveal hidden facets of host-circuit physiology. *PloS one*, *5*(7), e11909.

4. Danino, T., Mondragón-Palomino, O., Tsimring, L., & Hasty, J. (2010). A synchronized quorum of genetic clocks. *Nature*, *463*(7279), 326-330.

5. Stevens, A. M., & Greenberg, E. P. (1997). Quorum sensing in Vibrio fischeri: essential elements for activation of the luminescence genes. *Journal of Bacteriology*, *179*(2), 557-562.

6. Gottesman, S., Roche, E., Zhou, Y., & Sauer, R. T. (1998). The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes & development*, *12*(9), 1338-1347.

7. Wróbel, B., & Wę, G. (1998). Replication regulation of ColE1-like plasmids in amino acid-starvedEscherichia coli. *Plasmid*, *39*(1), 48-62.

8. Tomizawa, J. I. (1986). Control of ColE1 plasmid replication: binding of RNA I to RNA II and inhibition of primer formation. *Cell*, *47*(1), 89-97.

9. Panayotatos, N. (1984). DNA replication regulated by the priming promoter. *Nucleic acids research*, *12*(6), 2641-2648.

10. Selzer, G., Som, T., Itoh, T., & Tomizawa, J. I. (1983). The origin of replication of plasmid p15A and comparative studies on the nucleotide sequences around the origin of related plasmids. *Cell*, *32*(1), 119-129.

11. Stricker, J., Cookson, S., Bennett, M. R., Mather, W. H., Tsimring, L. S., & Hasty, J. (2008). A fast, robust and tunable synthetic gene oscillator. *Nature*, *456*(7221), 516-519.

12. Mather, W., Bennett, M. R., Hasty, J., & Tsimring, L. S. (2009). Delay-induced degrade-and-fire oscillations in small genetic circuits. *Physical review letters*, *102*(6), 068105.

13. Cookson, N. A., Mather, W. H., Danino, T., Mondragón-Palomino, O., Williams, R. J., Tsimring, L. S., & Hasty, J. (2011). Queueing up for enzymatic processing: correlated signaling through coupled degradation. *Molecular systems biology*, *7*(1), 561.

14. Elowitz, M. B., & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, *403*(6767), 335-338.

15. Gardner, T. S., Cantor, C. R., & Collins, J. J. (2000). Construction of a genetic toggle switch in Escherichia coli. *Nature*, *403*(6767), 339-342.

16. Roquet, N., & Lu, T. K. (2014). Digital and analog gene circuits for biotechnology. *Biotechnology journal*, *9*(5), 597-608.

17. Brophy, J. A., & Voigt, C. A. (2014). Principles of genetic circuit design. *Nature methods*, *11*(5), 508-520.

18. Rubens, J. R., Selvaggio, G., & Lu, T. K. (2016). Synthetic mixed-signal computation in living cells. *Nature communications*, *7*.

19. You, L., Cox, R. S., Weiss, R., & Arnold, F. H. (2004). Programmed population control by cell–cell communication and regulated killing. *Nature*, *428*(6985), 868-871.

20. Prindle, A., Samayoa, P., Razinkov, I., Danino, T., Tsimring, L. S., & Hasty, J. (2012). A sensing array of radically coupled genetic 'biopixels'. *Nature*, *481*(7379), 39-44.

21. Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H., & Weiss, R. (2005). A synthetic multicellular system for programmed pattern formation. *Nature*, *434*(7037), 1130-1134.

22. Moon, T. S., Lou, C., Tamsir, A., Stanton, B. C., & Voigt, C. A. (2012). Genetic programs constructed from layered logic gates in single cells. *Nature*, *491*(7423), 249-253.

23. Purcell, O., & Lu, T. K. (2014). Synthetic analog and digital circuits for cellular computation and memory. *Current opinion in biotechnology*, *29*, 146-155.

24. Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R., & Benenson, Y. (2011). Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science*, *333*(6047), 1307-1311.

25. Kobayashi, H., Kærn, M., Araki, M., Chung, K., Gardner, T. S., Cantor, C. R., & Collins, J. J. (2004). Programmable cells: interfacing natural and engineered gene networks. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(22), 8414-8419.

26. Friedland, A. E., Lu, T. K., Wang, X., Shi, D., Church, G., & Collins, J. J. (2009). Synthetic gene networks that count. *science*, *324*(5931), 1199-1202.

27. Tamsir, A., Tabor, J. J., & Voigt, C. A. (2011). Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. *Nature*, *469*(7329), 212-215.

28. Tabor, J. J., Salis, H. M., Simpson, Z. B., Chevalier, A. A., Levskaya, A., Marcotte, E. M., Voigt, C. A., & Ellington, A. D. (2009). A synthetic genetic edge detection program. *Cell*, *137*(7), 1272-1281.

29. Bonnet, J., Yin, P., Ortiz, M. E., Subsoontorn, P., & Endy, D. (2013). Amplifying genetic logic gates. *Science*, *340*(6132), 599-603.

30. Siuti, P., Yazbek, J., & Lu, T. K. (2013). Synthetic circuits integrating logic and memory in living cells. *Nature biotechnology*, *31*(5), 448-452.

31. Bonnet, J., Subsoontorn, P., & Endy, D. (2012). Rewritable digital data storage in live cells via engineered control of recombination directionality. *Proceedings of the National Academy of Sciences*, *109*(23), 8884-8889.

32. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., ... & Tinevez, J. Y. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods*, *9*(7), 676-682.

33. Škulj, M., Okršlar, V., Jalen, Š., Jevševar, S., Slanc, P., Štrukelj, B., & Menart, V. (2008). Improved determination of plasmid copy number using quantitative real-time PCR for monitoring fermentation processes. *Microbial cell factories*, *7*(1), 6.

34. Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., ... & Kummer, U. (2006). COPASI—a complex pathway simulator. *Bioinformatics*, *22*(24), 3067-3074.