UNIVERSITY OF CALIFORNIA SAN DIEGO

Highly multiplexed microfluidics for dynamic genome interrogation, synthetic gene circuit screening, and multi-target biosensing applications.

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

in

Bioengineering

by

Nicholas Sigmund Csicsery

Committee in charge:

Professor Jeff Hasty, Chair
Professor Nan Hao
Professor Prashant Mali
Professor Lorraine Pillus
Professor Yingxiao Wang

2019

The dissertation of Nicholas Sigmund Csicsery is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California San Diego

2019

DEDICATION

To Judy, Jackie, Sig, Kristine, and the rest of my loving family and friends.

EPIGRAPH

Everything was beautiful and nothing hurt.

—Kurt Vonnegut, Slaughterhouse-Five

You can always count on zntA.

—Old Bonnerian proverb

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Thank you to my expansive network of family, friends, colleagues, and advisors for immeasurable scientific and emotional support. Thank you to the members of the Biodynamics Lab, without whom many additional experiments would have gone awry, and even more unexplored. Thank you to BioCircuits Institute, the students and staff at Lafayette Elementary, and the many outreach volunteers who joined me in spreading science literacy throughout the San Diego area, Washington, D.C., and the Hawaiian archipelago. Finally, thank you to my advisor, Jeff Hasty, for enabling an enriching graduate experience and always encouraging me to research, teach, learn, and grow.

Chapter 2 contains material being prepared for submission as A custom high-throughput microfluidic and optical platform for the continuous culturing and measurement of *E. coli* and *S. cerevisiae*. Csicsery, Nicholas*, Stasiowski, Elizabeth*, Thouvenin, Gregoire, Graham, Garrett, O'Laughlin, Richard, Mather, William H., Ferry, Michael, Cookson, Scott, and Hasty, Jeff. (*equal contribution) The dissertation author was one of the primary authors and researchers of this material.

Chapter 3 contains material under review at the *Proceedings of the National Academy of Sciences* as Genome-scale transcriptional dynamics and environmental biosensing. Graham, Garrett*, Csicsery, Nicholas*, Stasiowski, Elizabeth*, Thouvenin, Gregoire*, Mather, William H., Ferry, Michael, Cookson, Scott, and Hasty, Jeff. (*equal contribution) The dissertation author was one of the primary authors and researchers of this material.

Chapter 4 contains material being prepared for submission as A multiplexed microfluidic platform for real-time heavy metal sensing and dynamic phenotype screening. Csicsery, Nicholas, Lezia, Andrew, O'Laughlin, Richard, Stasiowski, Elizabeth, Baumgart, Leo, and Hasty, Jeff. The dissertation author was the primary author and researcher of this material.

Chapter 4 also contains material being prepared for submission as A custom high-

throughput microfluidic and optical platform for the continuous culturing and measurement of *E. coli* and *S. cerevisiae*. Csicsery, Nicholas*, Stasiowski, Elizabeth*, Thouvenin, Gregoire, Graham, Garrett, O'Laughlin, Richard, Mather, William H., Ferry, Michael, Cookson, Scott, and Hasty, Jeff. (*equal contribution) The dissertation author was one of the primary authors and researchers of this material.

Chapter 4 also contains material under review at *Science Advances* as Interfacing gene circuits with microelectronics through engineered population dynamics. Din, M. Omar*, Martin, Aida*, Razinkov, Ivan, Csicsery, Nicholas, and Hasty, Jeff. (*equal contribution) The dissertation author was a secondary author and researcher of this material.

VITA

2013	B. S. in Biological Systems Engineering *summa cum laude*, University of California, Davis
2019	Ph. D. in Bioengineering, University of California, San Diego

PUBLICATIONS

Graham, Garrett*, Csicsery, Nicholas*, Stasiowski, Elizabeth*, Thouvenin, Gregoire*, Mather, William H., Ferry, Michael, Cookson, Scott, and Hasty, Jeff. "Genome-scale transcriptional dynamics and environmental biosensing", *Proceedings of the National Academy of Sciences*. (Under review) (*equal contribution)

Din, M. Omar*, Martin, Aida*, Razinkov, Ivan, Csicsery, Nicholas, and Hasty, Jeff. "Interfacing gene circuits with microelectronics through engineered population dynamics", *Science Advances*. (Under review) (*equal contribution)

ABSTRACT OF THE DISSERTATION

Highly multiplexed microfluidics for dynamic genome interrogation, synthetic gene circuit screening, and multi-target biosensing applications.

by

Nicholas Sigmund Csicsery

Doctor of Philosophy in Bioengineering

University of California San Diego, 2019

Professor Jeff Hasty, Chair

The rise of DNA sequencing and synthesis technologies over the past two decades has ushered in a new wave of forward engineering genetic circuits, synthetic biology. Synthetic biology has since been deployed for applications spanning therapeutics, industrial chemical biosynthesis, and environmental sensing. Coupled with advances in genomics and systems biology, synthetic biology has revolutionized our ability to investigate biological networks. Further utilized with the fine-tuned experimental control of microfluidics, synthetic biology has enabled the precise interrogation of single nodes in these biological networks. Recently, the emergence of genome-scale microfluidic devices has bridged the gap between the scale and throughput of

-omics technology and the dynamics achievable with microfluidics. Towards this end, we have developed an elegant, straightforward microfluidic platform capable of monitoring the temporal gene expression of 2,176 unique microbes with both research and industrial applications. In Chapter 1, I provide a brief overview of synthetic biology and microfluidics in the context of biological research. In Chapter 2, I describe the high-throughput microfluidic platform we have engineered and the protocol for building these devices. In Chapter 3, I demonstrate the platform's utility as an environmental biosensor and research tool, where the dynamics of 1,807 *E. coli* GFP-promoter strains coupled with machine learning algorithms are used to detect the presence of six heavy metals in real-time in both laboratory and real-world settings. Finally, in Chapter 4, I show the device's application for the dynamical screening of synthetic gene circuit libraries. In all, I highlight the need for the further development of such multiplexed microfluidic platforms and demonstrate their utility for biological research, synthetic gene circuit engineering, and environmental biosensing.

Chapter 1

Introduction

1.1 Synthetic biology and microfluidics: Tools for the study and forward engineering of biology

1.1.1 Towards forward engineering gene circuits

Synthetic biology is a rising and interdisciplinary field, drawing from the technology, knowledge, and technical approaches of engineering and the biological sciences to create novel biological systems. While genetic engineering and recombinant DNA technologies have existed since the 1970s [1], the early 2000s opened with a new wave of genetic modification based on the arrangement of genetic elements to produce novel cellular behavior [2–4]. This new iteration of genetic engineering, synthetic biology, was accelerated forward by technological leaps in the ability to both read and write DNA. Drawing from established engineering disciplines, synthetic biology seeks to enable the forward engineering of biological systems. With engineering principles, such as computational modeling and the standardization, decoupling, and abstraction of genetic elements [4], functional biological modules [5] have been created and interfaced together, including genetic toggle switches [2], oscillators [3, 6], and logic gates [7]. These

synthetic biology modules, or genetic circuits, have been employed for numerous applications over the past 20 years, thrusting synthetic biology to a market size of $12.80-13.96 billion in 2018 [8].

1.1.2 Applications of synthetic biology

Synthetic biology has addressed a wide range of application since its emergence in the early 2000s, including therapeutics [9], biomaterials [10], biofuels [11], pharmaceutical bioproduction [12], and biosensing [13]. However, biological complexity still exists and limits the true forward engineering of biological systems. To overcome this, researchers must acquire a deeper understand of basic biology, address the challenges of biological noise and evolution, and further characterize synthetic gene circuits. Fortunately, towards this end, synthetic biology often serves as the solution to its own obstacles, advancing biological research by enabling novel ways to investigate genetic networks in biological systems from the bottom-up. However, research and development in synthetic biology still remains as much as a scientific pursuit as it does an engineering discipline.

1.1.3 Interfacing microfluidics with biology

Microfluidics is a widespread technology for the precise control of fluids on the microliter and sub-microliter scales. Originating with the development of inkjet printers in the 1950s, microfluidics have since been deployed in the biological sciences. Notably, microfluidics have served as a useful tool for the miniaturization of established laboratory protocols, colloquially referred to as lab-on-a-chip devices [14], as well as the facilitation of novel biological assays. Beyond tools for basic research, a major area of commercial expansion for microfluidics is in the detection and reporting of compounds of interest from small sample sizes, particularly useful for point-of-care diagnostics or real-time monitoring of clinical or environmental samples [15].

Within synthetic biology, and in the Biodynamics Laboratory, microfluidics have served as a powerful tool for scientific discovery. Coupled with time-lapse fluorescence microscopy, microfluidic platforms provide precise imaging and environmental control to investigate the underlying dynamics of biological networks in microorganisms [16]. In our lab, this has enabled the development of genetic oscillators [6, 17, 18], studies of cellular aging or competitive fitness in dynamic environments [19, 20], and clinically relevant cancer therapies [9]. These developments have spanned micrometer scales single cell analysis [19] to the millimeter scale, observing 2.5 million cells across a 5 mm device [18].

Conventional laboratory microfluidic devices consist of an elastomer polydimethylsiloxane (PDMS) monolith bonded to glass. The PDMS is cast from a photoresist patterned silicon wafer made using standard photolithography. After surface treatment, the PDMS and glass are bonded together and cured at high temperatures (80°C) for several hours. Cells are then loaded onto devices after bonding, often requiring a distinct waste port for each unique strain to avoid cross-contamination [16]. Due to space limitations and experimental complications, conventional microfluidic devices are generally low-throughput, allowing users to track the behavior of a few nodes to answer narrow biological questions [21–23].

A handful of studies have demonstrated how microfluidic parallelization and automation allows for the simultaneous tracking of hundreds to thousands of strains from available fluorescent libraries for *S. cerevisiae* [24–26] and *E. coli* [27, 28], utilizing the precise control of microfluidics to generate -omics scale data sets. While significantly expanding the utility of a single device, highly parallelized microfluidics typically consist of multiple layers with pneumatic connections allowing valves to control media flow, cell loading, and chemical delivery [24,29,30]. This leads to additional experimental complexity compared to single-strain microfluidics, which can complicate experimental setup, limit control of environmental dynamics, and shorten experimental run-time. Thus, while incredibly useful to a trained technician, these obstacles impede the widespread adaptability of highly-multiplexed microfluidics throughout academic and industrial labs.

In the following chapters, I describe the development and application of a broadly applicable highly-multiplexed microfluidic platform. This single-layer, single-inlet, single-outlet device exhibits microfluidic and experimental simplicity while enabling the real-time monitoring and environmental control of over 2,000 *E. coli* or *S. cerevisiae* strains over week timescales. I demonstrate the utility of this platform to basic biology research, enabling the generation of -omics scale gene expression data at a 3 minute temporal resolution. I show the platform's potential to assist synthetic biology design as a tool for the high throughput screening of variant gene circuit constructs. Further, I describe the immediate real-world application of this platform as an environmental biosensor for the real-time detection of heavy metals in water supplies.

Chapter 2

Highly-multiplexed microfluidics for continuous culturing and monitoring of *E. coli* and *S. cerevisiae* strain libraries

2.1 Introduction

Genome-scale technologies have transformed our understanding of the biomolecular signaling networks that underpin gene function, cellular behavior, and drug responses. Omics-level analysis has cemented the view that biological signal processing is not the result of linear pathways, but an emergent property of complex networks whose functions and dynamics we now seek to understand [18, 31–35]. In model organisms, signaling networks are often elucidated by studying the changing patterns of gene expression in reaction to experimentally-induced environmental perturbations [2, 31, 32]. However, the high-throughput, genome-wide techniques (including RNA-seq [36, 37], ribosome profiling [38], mass spectrometry [39], and microarray-based expression profiling [40, 41]) that have successfully been applied to this problem are destructive in nature and only offer snapshots of a cell's state [42]. Technologies that acquire

single time point data do not capture the information encoded in the dynamics of biomolecular networks [43], which are complex, time-dependent signals [44–46].

Over the past two decades, microfluidic devices coupled to time-lapse fluorescence microscopy have evolved to bridge the gap of -omics data and precisely controlled dynamic environments, resulting in high-quality time series data for thousands of signaling network constituents in live cells [47,48] for a variety of model organisms and cell lines [24–28]. Here, we present an advance in microfluidics with the pipeline for the simultaneous loading, culture, and measurement of over 2,000 unique *E. coli* or *S. cerevisiae* strains on a straightforward, single-inlet, single-outlet, and single-layer device.

2.1.1 Existing multiplexed microfluidic platforms

Some of the first high-throughput microfluidic devices for cellular analysis were published from the Quake lab at Stanford University. In 2002, the group published a device consisting of 256 individually addressable chambers controlled by 2,056 microvalves where two separate substrates can be loaded into a compartmentalized chamber, mixed pairwise, and independently retrieved [49]. This device was used as a microfluidic comparator to successfully screen and retrieve *E. coli* cells expressing a cytochrome c peroxidase enzyme from a heterogeneous population. A similar platform was published in 2007 for mammalian cell culture, with 96 individually addressable cell chambers [50]. Though this device can only load a maximum of two unique strains, it introduced multiplexed valve systems to microfluidics and paved the way for many valve-based devices. This device was used to investigate the influence of transient stimulation schedules on human primary mesenchymal stem cells. While these devices primarily showcased the ability to address a single strain with under multiple conditions, they introduced multiplexed valve systems to microfluidics and paved the way for many valve-based devices.

The Maerkl lab at the Swiss Federal Institute of Technology in Lausanne (EPFL) further developed valve-based microfluidic technology to develop a 1,152-chamber microfluidic

device for monitoring *S. cerevisiae* at the single cell level [24]. Cells are first spotted onto an epoxysilane-coated coverslip using a DNA microarrayer. The multi-layer PDMS device is then aligned and bonded to the coverslip. The device has four valves that are opened and closed in a specific sequence, controlling the direction of flow to prime the chip with media, prevent cross-contamination, and allow the cells to continuously grow as a monolayer in each chamber. This device is arrayed into 3, 384-chamber sections that are fluidically isolated from each other and allow a medium switch between two sources. The temporal resolution for imaging this device is twenty minutes. The Maerkl group measured gene expression over six hours of growth and seven hours of DNA damage exposure. While this device is the most high-throughput device currently available for *S. cerevisiae* monitoring, the temporal resolution is not sufficient for measuring dynamics in all signaling networks in *S. cerevisiae*. For example, in the oxidative response network, the time scale of protein expression change is on the order of minutes [51]. Additionally, this device only monitors strains for less than a day, preventing the monitoring of longer dynamics such as fermentation and shifts in metabolism and growth phases. Finally, this platform only allows for one medium shift and does not allow for a dynamic input, despite such control being one of the most salient advantages of microfluidics [16].

The Maerkl lab additionally developed a valve-based microfluidic device with 768 chambers for monitoring *E. coli* [29]. The same spotting technique is used as their *S. cerevisiae* device, where cells are arrayed onto a coverslip using a DNA array spotter and a multi-layer PDMS device is aligned and bonded to the coverslip. While this device can culture and monitor strains for up to one week, it requires a complex sampling process where a series of valves are opened and closed in order control the flow of medium, sample, and lysis buffer every hour throughout the duration of the experiment. During this sampling process, the sample is flowed onto the device for 10 minutes, the cells are then exposed to the sample and media via diffusion for 45 minutes, and then lysis buffer is flown on to clear the channels for ten minutes. This discontinuous exposure to a fresh sample could lead to a change in the microenvironment in each biopixel over time due to

nutrient metabolism and waste secretion by the cells. Additionally, this sampling step inhibits the dynamic input of the sample. This device was used to characterize strains and as a biosensor to detect varying levels of arsenic using a custom optical system.

Finally, a 1,500 chamber mammalian cell culture device was developed by the Tay Lab at the University of Chicago and is a built to culture single cells, two-dimensional populations, or 3D neurospheres of stem cells [30]. Each chamber can have independent culture conditions including cell type and density, and can also have a unique set of signaling molecules, growth factors, or drugs delivered to the chamber. The device has an on-chip chemical formulator, peristaltic pump, and valve system to mix various chemicals and deliver them to individual chambers with sub-minute temporal resolution. Contrary to the previous valve-based devices where cells are spotted prior to bonding the PDMS device to the glass slide, this device is first bonded and then cells are loaded onto the PDMS-glass device through a semi-automatic loading program through the channels and on-chip valve system. A typical week long experiment results in the tracking of 30,000 individual cells with 1,500 dynamic individual conditions. While this device is capable of 1,500 dynamic inputs of chemicals, it is not used in the study to look at many different cell types in one experiment. This is potentially due to the requirement of loading each strain individually using the valve system.

While all three of the devices described above have many strengths with respect to monitoring the gene expression of their organism of interest, all share the disadvantage that these multi-layer, valve-based microfluidic devices have complex experimental protocols that are cumbersome to set-up and require hardware that limits the devices' usability in an industrial setting. These shortcomings and those listed above impede the widespread adoption of a potentially transformative technology and its application towards signaling pathway dynamics and associated biomedical applications [25]. To this end, by sacrificing the ability to individually address each cell culture chamber, we have developed a novel single-layer, two-fluidic-port device for the high-throughput continuous culture and measurement of over 2,000 *E. coli* or *S. cerevisiae* over

week timescales. This chapter describes the optimization performed to fabricate, load, and bond these devices and concludes with a full detailed protocol for their manufacturing.

2.1.2 Fluorescent strain libraries

To validate the use of our multiplexed microfluidic platform we tested the viability, long-term growth, and fluorescent dynamics of two strain libraries: an *E. coli* promoter library from the Alon group [27] and an *S. cerevisiae* library from the O'Shea and Weissman groups [26]. The *E. coli* is composed of 1930 unique 50-150 base pair promoter regions from the MG1655 *E. coli* genome inserted into a low copy, pSC101, plasmid upstream of a strong ribosome binding site and GFP (Figure 2.1A). The *S. cerevisiae* library is comprised of 4,159 unique strains in which GFP fusion proteins were integrated into the BY4741 yeast genome, to be expressed by endogenous chromosomal promoters (Figure 2.1B). Upon their creation, both libraries were used to measure gene expression and promoter activity in their respective studies, though only under batch growth conditions [26, 27, 52]. In our effort to characterize these library dynamics under continuous growth conditions, we sought to engineer a multiplexed microfluidic device that could simultaneously culture over 2,000 unique strains, amounting to the entire *E. coli* library or half the *S. cerevisiae* library on a single chip.

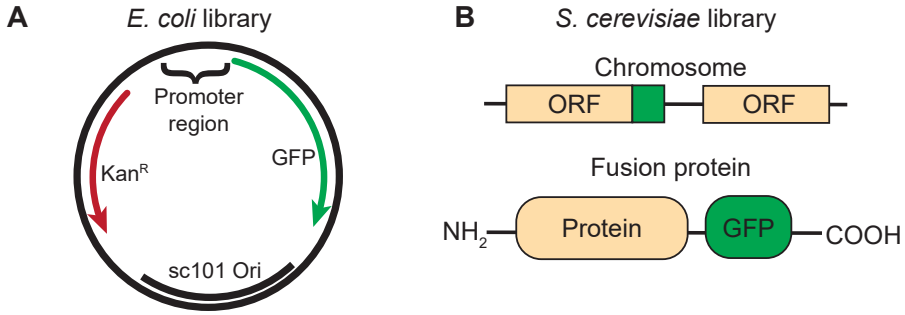

Figure 2.1: Strain library architecture. (A) *E. coli* library members are transformed with a plasmid, each with a unique 50-150 base pair promoter region upstream of GFP. (B) *S. cerevisiae* library members have GFP genomically integrated, each with a unique native-protein GFP fusion.

2.2 Microfluidic device design

2.2.1 Cell loading requirements

Traditional loading of microbial cell cultures onto microfluidics requires an individual fluidic line and connecting port into the microfluidic assembly for each unique strain in order to minimize the risk of strain-cross contamination [16]. The use of multi-layer devices with microvalve arrays can offer a workaround to excessive fluidic connections, but introduces its own complexities and an increased risk of cross contamination. To bypass superfluous fluidic connections while maintaining a simple and easily adoptable fluidic architecture we found it necessary to develop a method to load strains directly onto the device before bonding PDMS to glass (Figure 2.2). As an additional constraint, loading 2,000 strains onto either the glass or PDMS components of a microfluidic device would need to be rapidly performed, as PDMS-glass bond strength weakens with time between plasma exposure and component bonding.

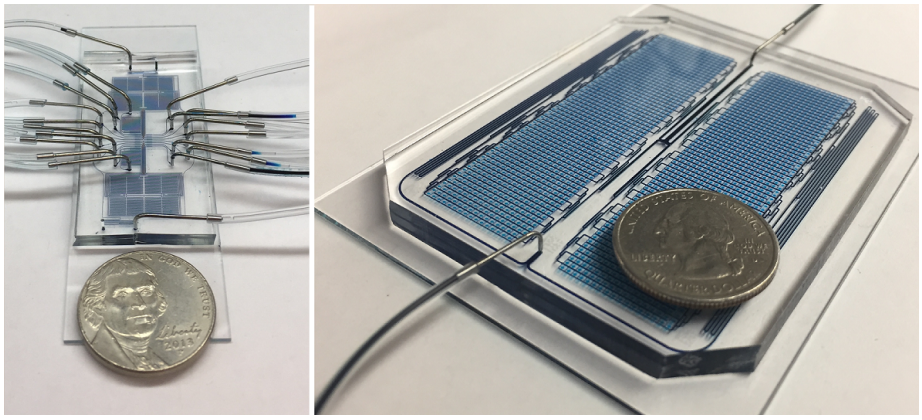

Figure 2.2: Device complexity in multiplexed microfluidics. An early stage 18-strain microfluidic device requiring 20 fluidic connections (left) compared to the final 2,176-strain device with 2 fluidic connections (right). The 18-strain device is loaded using conventional fluidic methods while the 2,176-strain device is loaded using our novel spotting protocol.

There is no off-the-shelf equipment that can load tens to thousands of strains onto a microfluidic device. To meet this need, we adapted the Singer ROTOR spotting robot from its original purpose as a tool to transfer cells between liquid well-plates or agar plates using contact

(agar) and surface tension (liquid), to a tool that could transfer cells from agar to a microfluidic device. The microfluidic design requirements were constrained to the capabilities of this robot. The requirements included the following: a cell trap region at least 400 μm in diameter where each initial cell spot would be placed, 1.125 mm spacing between each cell trap in both the x and the y dimension due to SBS-format limitations, and a region for exponential cells to grow and have sufficient diffusion of nutrients where the measurements would be taken.

2.2.2 Cell trap design

Before scaling up to the final 2,000-strain microfluidic device, a 32-strain mid-scale device was created to verify our spotting protocol (Figure 2.3). This mid-scale device features a 4x8 array of cell spotting regions, each connected to 10 cell chambers as seen in Figure 2.3c. Spacing between the spotting regions is 1.125 mm for compatibility with standard SBS-format 6144-density spacing. Flow through the device is such that each cell chamber region is fluidically isolated to prevent strain cross-contamination. A single strain bank is detailed in Figure 2.3a. Each cell spotting region (a1) can be loaded with a different cell strain. Upon wetting the device, cells grow along the feeder channels (a2) and fill the cell chamber (a3). The cells grow on fresh media delivered by the media channel (a4). A linker channel (a5) allows cross-seeding of the cell chambers in order to continually refill any chambers whose cells are washed out.

The geometry of individual cell chambers is shown in Figure 2.3b. The cell chambers are designed 10 μm wide to minimize internal convective flow from the media channel, improving cell retention. They are fabricated as 50 μm tall and closely spaced at 30 μm in order to maximize the cellular fluorescence signal integrated along the orthogonal optical imaging path. The full chip layout is shown in Figure 2.3c and an image of a single strain bank in exponential growth with E. coli in Figure 2.3d.

While the cell trap in Figure 2.3 was used for development of the spotting protocol, several trap geometries were tested for the final design. Trap designs were tested in order to maximize

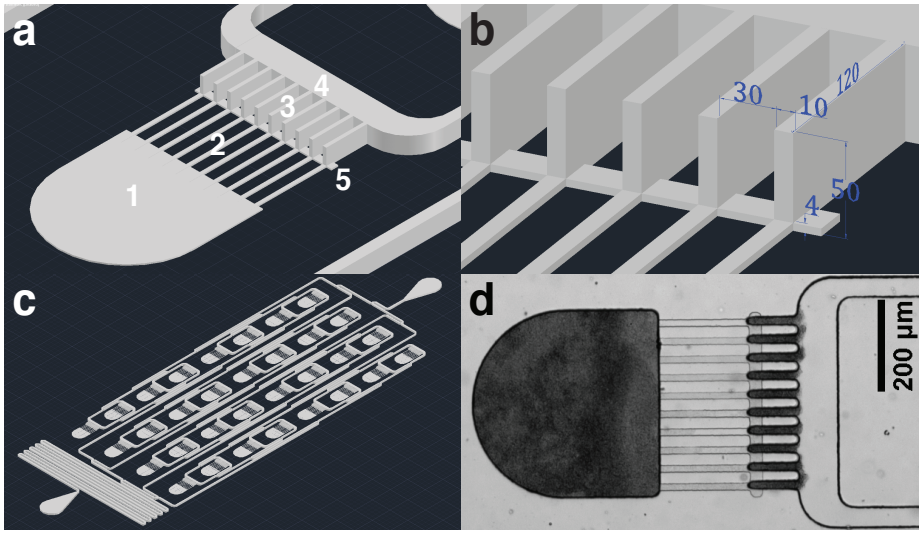

Figure 2.3: Geometry of the mid-scale 32-strain microfluidic device. a) Schematic of single strain bank, with a cell spotting region (1), feeder channels (2), cell traps (3), media channel (4), and linker channel (5). b) Detailed dimensions (µm) of the cell traps, optimized for cell growth, retention, and optical signal. c) Overview of the mid-scale device with 32-strain banks connected in parallel. d) Microscope image of *E. coli* growing exponentially in one strain bank.

E. coli cell retention, homogeneous growth states, and fluorescence signal while minimizing excess cell mass that would lead to clogging and therefore a shorter device lifetime (Figure 2.4). Figure 2.4a shows the cell traps initially designed and described above, referred to informally as "gills". Gill traps had a high aspect ratio which resulted in maximized fluorescence signal, however cell retention was poor, due to convective flow into the first gill and flowing out of the tenth gill. The Gill-Shunt design (Figure 2.4b) was designed to reduce convective flow in the gill traps by providing a path of lesser resistance for flow to follow before cells filled. Although this design helped redirect flow, cell retention was still variable. Furthermore, the high aspect ratio of the gill traps made devices difficult to fabricate with techniques other than PDMS lithography, limiting the possibility of more commercially viable fabrication methods such as hot embossing. The cell traps in Figure 2.4c were designed to have a dead volume region in order to better retain cells. While cell retention improved, cell growth states were variable, and resulted in variable fluorescence responses. Figure 2.4d shows the simplest trap possible, a single open side trap in which cells are directly spotted. Cells in this 50 µm tall trap accumulated too much biomass and

clogged the device quickly. Figure 2.4e shows a design similar to 2.4d with an added spotting post to constrict some biomass while also providing a support pillar for the spotting process. This design also resulted in excessive biomass. Figure 2.4f shows a funneled cone design also 50 μm in height that restricted cell mass. This design was iterated for use in the channel-free design (Figure 2.4g-j). Figure 2.4g has a 4 μm spotting region and 50 μm funneled cone. Figure 2.4h shows a teardrop wall in order to prevent upstream cells from gathering on downstream trap walls. It has a 4 μm spotting region and 50 μm funneled cone similar to Figure 2.4g. The narrow funneled cone prevented clogging and large biomass, but the 50 μm height still resulted in variability in the growth state. Similarly, Figure 2.4i also has a teardrop shape and 4 μm spotting region. Instead of a 50 μm cone, this design has 50 μm gills as the cell trap. The gills reduced cell mass and clogging, but still showed variability in growth state. Figure 2.4j has a 4 μm spotting region and 4 μm cone. This design limited biofilm formation, decreased growth state variability, and had sufficient fluorescence signal; however, this design was not compatible with the manifold channel design. This trap design was adapted to work with the manifold channel design (Figure 2.4k), successfully retaining cells in a homogeneous growth state with detectable fluorescence signal. The design in Figure 2.4k, referred to as "bulbs", was had the best overall performance and was the primary trap design used for the work presented in Chapter 3.

2.2.3 Channel system design

Traditionally in microfluidics, a binary splitting channel system is used, where each cell trap has an equal channel path length in order to have identical flow rates and fresh nutrients. However, as the number of desired cell traps increases, this channel system requires large amounts of space (Figure 2.5a). Moreover, the total number of cell traps must be a factor of 2^n, where n is the number of splitting events, in order to have even flow across all traps. Three additional channel systems were tested in order to decrease the feature footprint between traps, while maintaining sufficient nutrient delivery, preventing excessive buildup of cell mass, and preventing

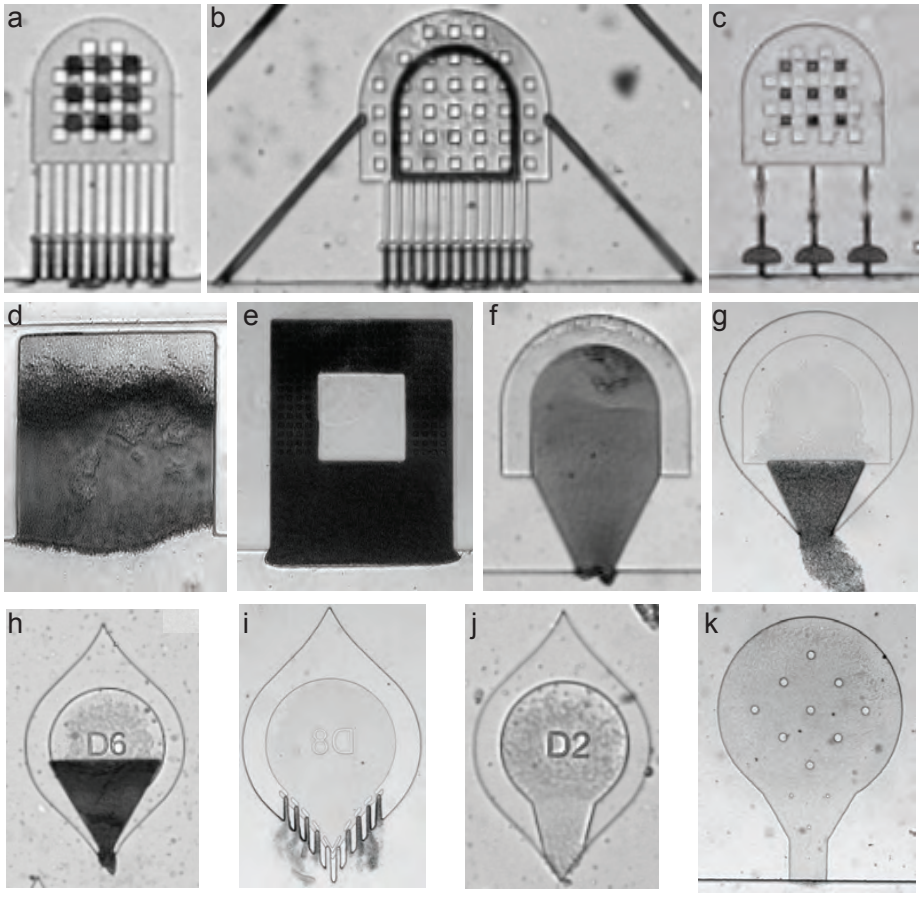

Figure 2.4: Cell trap designs tested to maximize *E. coli* cell retention, homogeneous growth states, and fluorescence signal. Gill traps with (a) and without (b) shunts inconsistently filled traps. Dead end gills (c) and side traps (d,e) lead to heterogeneous growth states. Channel-free traps (f-j) were tested with 4 μm (f,j) and 50 μm (g-i) openings with various opening sizes for cell retention. k) The final trap design used with the manifold channel system for Dynomics data collection in Chapter 3.

cross contamination (Figure 2.5b-d).

The shared channel system (Figure 2.5b) consists of a single inlet channel splitting into multiple channels of equal length. Each channel has many traps in series, with media from upstream traps flowing directly to downstream traps. This design ensures an even flow rate across all traps in a shared channel, regardless of clogging of fabrication defects. However, we found that downstream cell traps are cross-contaminated with effluent cells from upstream traps. The shared channel system is only successful if replicates of strains are loaded in a single channel,

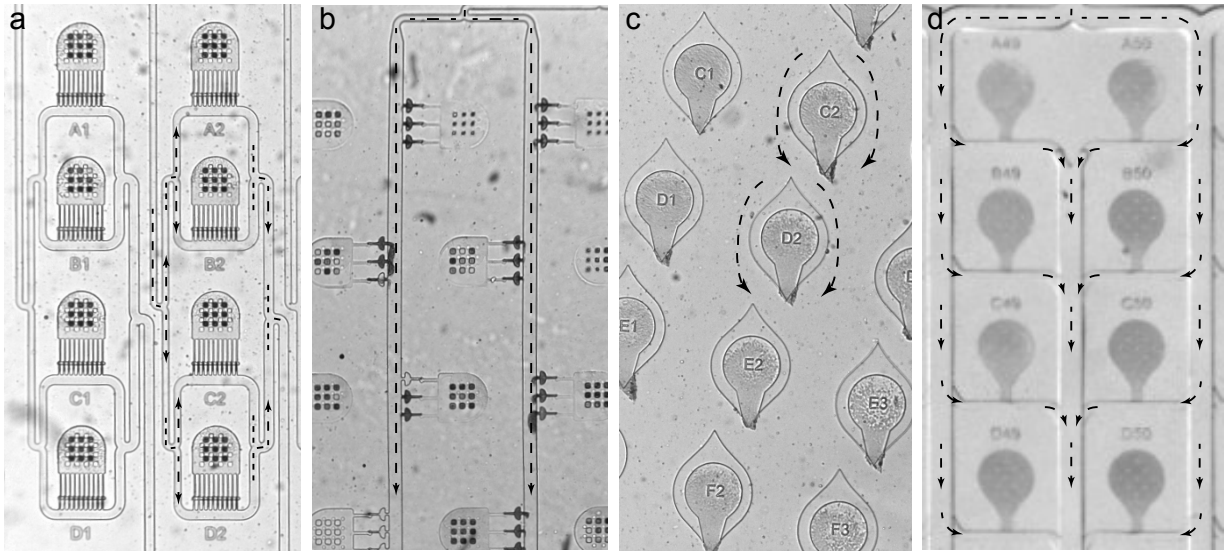

Figure 2.5: Channel systems tested for increasing the throughput of microfluidic devices constrained to 1.125 mm spacing between traps. a) Binary splitting channel system traditionally used in microfluidics. b) Shared channel system. c) Channel-free system. d) Manifold channel system which is used in the 2,176-strain device. Black arrows denote flow direction.

reducing the throughput of the device.

The channel-free system (Figure 2.5c) consists of an inlet channel feeding media into one large open area with media flowing around the top of cell traps. The traps were rotated 15° so that the laminar flow paths from the mouth of one cell trap would not directly feed another, reducing cross contamination, and maximizing nutrient diffusion. While the channel-free system has the smallest feature footprint, analysis using fluorescent microbeads showed that flow direction unexpectedly changes near some traps, resulting in cross-contamination. The unpredicted changes in direction of flow were likely due to small differences in the height of the open area due to imperfections in the silicon wafer fabrication process.

The manifold channel system (Figure 2.5d) consists of an inlet channel that splits into several vertical major media channels with a height of 230 μm. Major channels split into minor channels in parallel that are 50 μm tall. Each minor channel feeds one cell trap. The minor channels then combine again into shared outlet media channels. The manifold channel system meets the design requirements best, as cells see fresh media at each trap eliminating cross

contamination, and the channel footprint is small enough to fit between the required 1.125 mm spacing between cell traps. While the number of manifold columns is restricted to a factor of 2^n, any number of rows can be used. The manifold channel system results in about a 20% difference in flow rate down a column of cell traps; however, that did not affect the response of the *E. coli* strains tested during channel optimization. The manifold channel system was implemented in the large-scale device.

2.2.4 Device scale-up

With several trap and channel design vetted, we had the framework to scale the device up to the thousand-strain scale. The manifold channel system was scaled-up to include 2,176 strains as seen in Figure 2.6. The strain banks are arranged into two halves of 17x64 strain banks. A single inlet splits to feed media to each half of the device, with one outlet combining the flow to remove excess cells and media waste. With the manifold system, a variety of cell traps can be substituted based on experimental requirements. The gill and bulb designs are the two cell trap designs used most frequently in this work.

2.3 Device loading

2.3.1 Strain library handling and replication

To prepare strains of *E. coli* or *S. cerevisiae* for loading onto a microfluidic device, they are rearranged from their original 96-well density format. The strains are handled using the Singer ROTOR (Figure 2.7a). Bulk strain movement uses Singer 96-, 384-, 1536-, or 6144-density RePads (Figure 2.7b) whereas single colonies can be rearranged using the Singer Stinger (Figure 2.7c). Strains are maintained either in liquid cultures on 96- or 384-well liquid plates or in solid form on agar plates with a colony density of up to 6144 unique colonies per plate. To ensure that

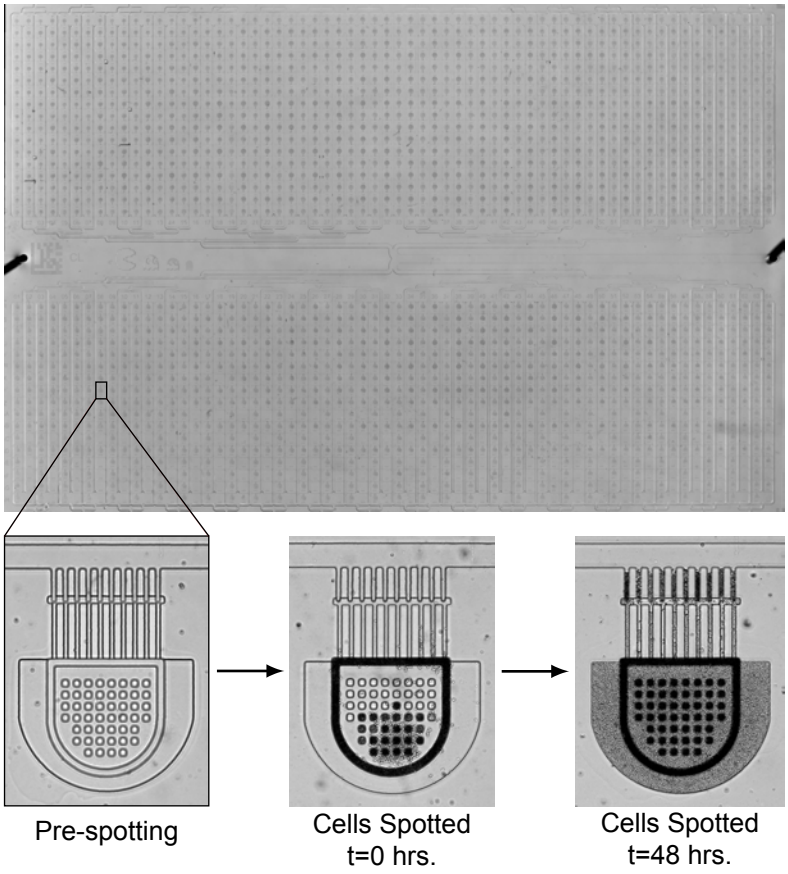

Figure 2.6: A large-scale, 2,176 strain microfluidic device. Image of the large-scale 2,176-strain device loaded with *E. coli*. The inset shows a single reservoir before and after cell spotting and revival.

colonies on agar plates do not cross-contaminate, growth times are limited such that colonies do not touch and are not too large for the RePad pins. We found that growth times for various colony densities were consistent with those previously reported [53].

The workflow for rearranging cells is shown in Figure 2.7d. First, 96-well liquid plates are spotted to 96-colony-density agar plates using the ROTOR. Next, the Stinger re-arrays strains onto four fresh agar plates at a 1536-density based on the desired arrangement on the microfluidic device. The four 1536-density agar plates are then combined to form a single 6144-density plate. 6144-density RePads are used to transfer cells from the 6144-density agar plate to the microfluidic device(s).

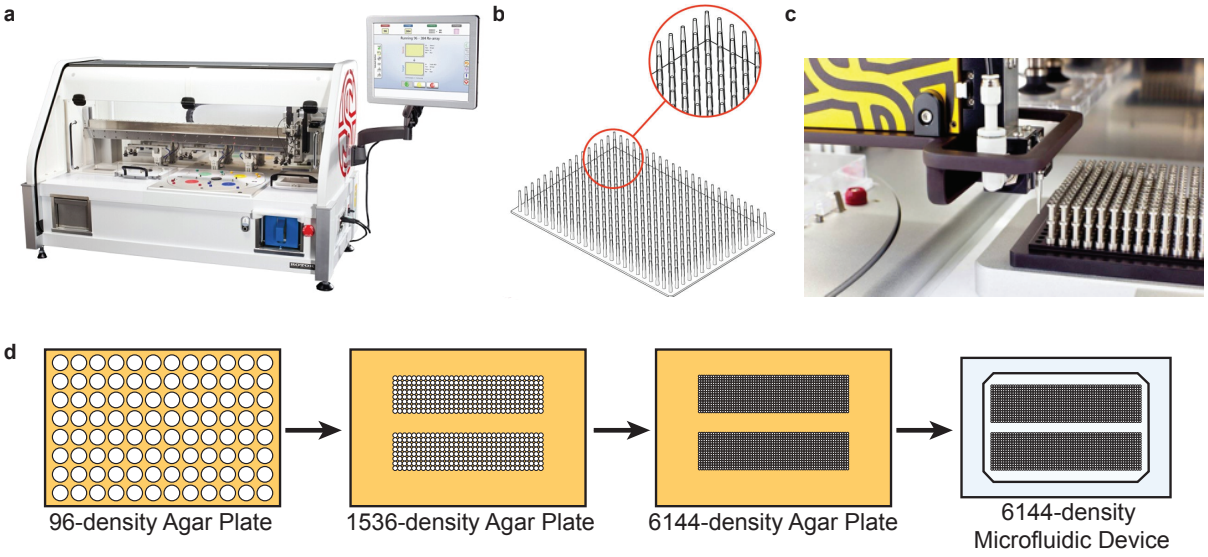

Figure 2.7: Cell arraying equipment and workflow. a) Singer ROTOR HDA. b) A 384-density long-pin RePad, used for bulk movement of strains. c) Singer Stinger, an attachment used for single colony movement. d) The workflow for loading strains onto a large-scale microfluidic device. Cells are re-arrayed from 96-density to 1536-density using the Singer Stinger. 1536-density plates are scaled up to a 6144-density and then spotted to the microfluidic device.

2.3.2 Robot loading optimization - PDMS unbonding

To determine the best cell loading method with the Singer ROTOR, both spotting to glass slides and to PDMS was explored. Initial device loading tests involved spotting cells from agar plates onto glass slides using the mid-scale device. PDMS and glass slides were cleaned and then exposed to oxygen plasma. After plasma exposure, glass slides were placed on a 3D printed insert compatible with the Singer ROTOR. Cells were spotted from the 6144-density agar plate onto the glass slides. Finally, PDMS was aligned to the cell array on the glass slides using a photolithography mask aligner, pressed together, and bonded at 37°C for two hours. In all, time between plasma exposure and glass-PDMS bonding took 20-40 minutes.

Spotting to glass resulted in poor bonding between the glass and PDMS. Regions surrounding spotted cells often unbonded, causing clogging and cross-contamination (Figure 2.8a). We hypothesized that the increased hydrophilicity of the glass slides after oxygen plasma treatment was causing a thin layer of liquid to spread around the cells and prevent bonding. To test this

hypothesis, cells were grown on agar plates containing 10 μM fluorescein and spotted to both plasma treated and untreated glass slides. Fluorescence imaging of glass slides showed that significantly more liquid spreading occurred on plasma treated glass slides. Adding a five second delay between picking up cells and spotting them allowed cells to dry and resulted in less liquid spreading, but also led to inconsistent spotting and smaller cell spots (Figure 2.8b).

To spot cells to PDMS, cells were first transferred from a 6144-density plate to an acrylic insert compatible with the Singer ROTOR to establish alignment markers. Next, the PDMS device was aligned to the cell spots on the acrylic insert using a photolithography mask aligner. After alignment, we exposed glass slides and our PDMS aligned to the acrylic insert to oxygen plasma. We then spotted cells to the PDMS using the Singer ROTOR, pressed together the PDMS and glass slide, and bonded the device at 37°C for 2 hours. For this method, time between plasma exposure and glass-PDMS bonding is about 10 minutes.

Experiments where cells were spotted to PDMS resulted in better bonding than those spotted to glass slides, but the hydrophilicity of the PDMS caused moisture to wick to the edges of the reservoirs, resulting in some unbonding. This was solved by adding features to the reservoirs to increase surface area (Figure 2.8c). Additionally, spotting parameters on the Singer ROTOR including pin pressure, speed, offset, and overshoot were fine-tuned for spotting the optimal number of cells. Due to decreased process time and higher bonding efficiency, we chose spotting cells to PDMS as our primary loading method.

2.3.3 Robot loading optimization - cell viability

Cell health during experiments is essential; therefore, we optimized growth conditions before and after spotting cells to the microfluidic device to ensure that cells survived the spotting process and recovered during the experiment.

During early *S. cerevisiae* experiments, anywhere from 25-97% of spotted reservoirs per chip failed to grow (Figure 2.9). We hypothesized that this was due to cells drying out between

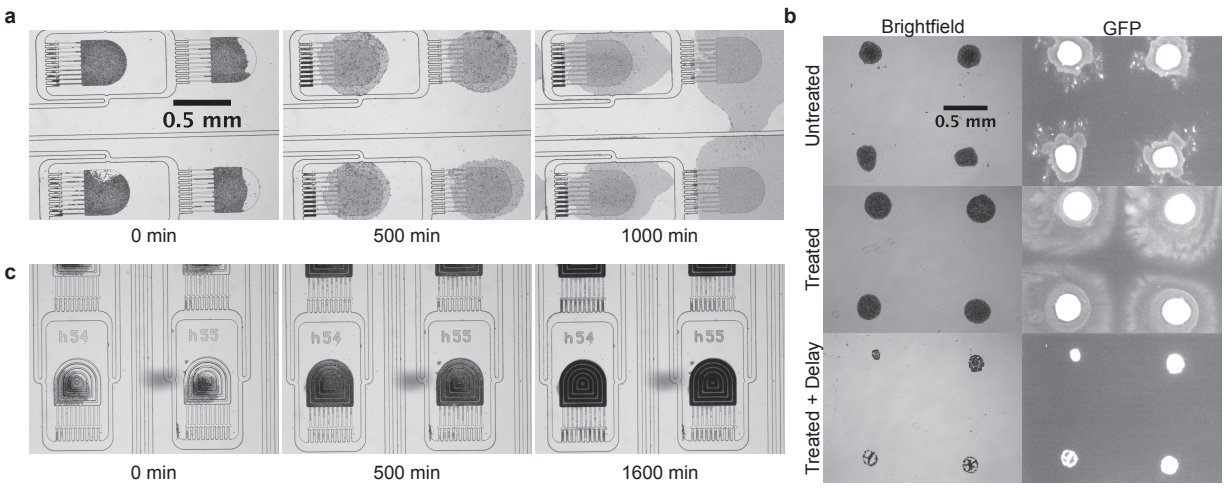

Figure 2.8: Minimizing chip unbonding after spotting onto microfluidic devices. a) Experiments with cells spotted to glass slides resulted in significant unbonding. b) Cells grown on 10 μM fluorescein agar plates that were either treated or untreated with oxygen plasma. Spreading of fluorescent fluid was largest for treated slides. Adding a 5-second delay prevented spreading but resulted in few cells. c) Cells spotted to PDMS patterned reservoirs resulted in well bonded devices.

being loaded onto the device and being wetted with media. We verified this using a yeast viability kit [54]. Cells growing on agar plates were spotted onto glass slides and then allowed to dry for variable periods of time. After drying, cells were stained with green- and red-fluorescent dyes, in which the red dye selectively stained nonviable cells. Using flow cytometry, we found fresh cells growing on agar plates to be 98% viable, whereas cells spotted to glass slides and immediately stained for viability were only 6% viable (Figure 2.9A).

To improve *S. cerevisiae* viability, we sought to increase the uptake and production of a cryoprotectant, trehalose, in the cells. Trehalose has been shown to increase the tolerance of *S. cerevisiae* to the osmotic stress associated with freezing and drying [55,56]. To increase trehalose uptake, cells were grown on yeast extract peptone (YP) plates supplemented with 2% Trehalose and 1% Galactose [57]. In microfluidic runs loaded from cells grown on trehalose plates, we observed complete cell growth. To increase trehalose production, we grew *S. cerevisiae* under heat shock conditions at 37°C. Cells grown by this method also exhibited complete recovery in microfluidic experiments. Cells grown on trehalose plates with heat shock did not demonstrate a

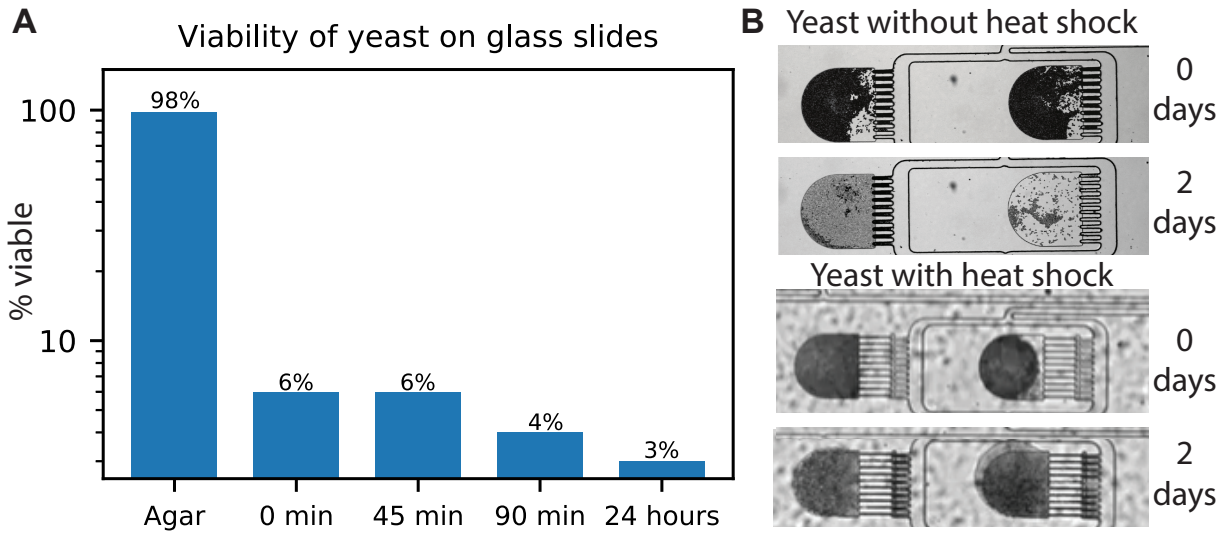

Figure 2.9: Optimizing yeast viability after contact transfer. (A) Cell viability after spotting onto glass slides. (B) Microfluidic experiments with and without heat shock to induce trehalose production in *S. cerevisiae*.

significant advantage over cells only grown on trehalose or only grown under heat shock, thus all further yeast runs were conducted with a simple heat shock. Future runs were conducted by growing yeast under heat shock conditions before spotting (Figure 2.9B).

Early *E. coli* experiments had complete revival on microfluidic devices when grown on LB+Kan agar plates. However, both *E. coli* and *S. cerevisiae* viability decreased after replicating the same set of strains for several months. Consequently, we thaw fresh overnight cells stored at -80°C every six months and repeat the preparation pipeline.

After *S. cerevisiae* and *E. coli* are spotted, they are grown on rich SC-His and LB+Kan media, respectively, for optimal growth. Once reservoirs and traps are filled, *S. cerevisiae* experiments are switched to a minimal media, YNB+Met+Ura+Leu, to decrease the fluorescence of media for an increased signal to noise ratio. Similarly, *E. coli* experiments are switched to a minimal media optimized for microfluidic experiments, HM9. Table 2.1 summarizes optimal growth conditions for *E. coli* and *S. cerevisiae*.

Table 2.1: Growth conditions for *E. coli* and *S. cerevisiae* for optimal device performance.

	Agar growth	**Microfluidics - grow up**	**Microfluidics - imaging**
E. coli	LB+Kan at 37°C	LB+Kan at 37°C	HM9+Kan at 30°C
S. cerevisiae	SC -His at 37°C	SC -His at 30°C	YNB+Met+Ura+Leu at 30°C

2.4 Device validation

After successfully developing a protocol for the loading, growth, and imaging of strains of both *E. coli* and *S. cerevisiae*, we verified that we could achieve continuous monitoring of fluorescence levels in library strains. For *E. coli*, 7 strains were arrayed evenly across the large-scale chip: the cadmium responsive *zntA*, the copper responsive *cueO*, the promoterless control U139, and three sigma factors, *rpoH*, *rpoS*, and *rpoD*. Baseline fluorescence expression of these promoters was not significantly different based on device position. Similarly, upon induction with cadmium or copper, the corresponding strains responded as expected, with *zntA* responding to cadmium and *cueO* responding to copper (Figure 2.10). Variation observed within these responses was not dependent on device position.

Similarly, *S. cerevisiae* was validated by running the entire library across two chips, and ensuring that we could measure gene expression changes in response to osmotic stress. The library was subjected to multiple short, 2-3 hour, pulses of hyperosmotic stress (1 M sorbitol) [58]. Three representative dynamic fluorescent responses were extracted and are shown in Figure 2.11b. While we observed *S. cerevisiae* to have higher auto-fluorescence than *E. coli*, it was still possible to measure changes in gene expression by fluorescence measurement.

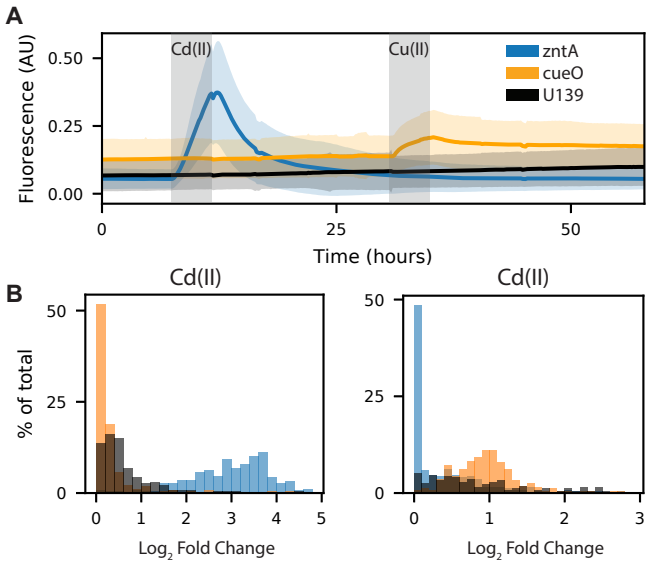

Figure 2.10: Library member response to cadmium and copper inductions. (A) Roughly 300 copies of 7 control strains were arrayed evenly across the device. Time series responses to cadmium and copper inductions are shown for zntA, cueO, and the promoterless U139 strain. Solid line represents mean, while shaded region represents standard deviation. (B) Histograms showing the percentage of each strain responding with a specific fold change to each metal induction.

2.5 Protocol for the contact loading of PDMS microfluidic devices with cells grown on agar plates.

Below is the step by step loading and bonding protocol for mid-scale and large-scale devices, for both *E. coli* and *S. cerevisiae*.

2.5.1 Wafer fabrication

Our group has previously described the steps to design and build a silicon wafer patterned with the features of the microfluidic device [16]. Briefly, the device is designed using AutoCAD or similar software, with each set of features of one desired height designed on an individual mask. Masks are printed by CAD/Art Services, Bandon, OR. A silicon wafer is built using standard soft photolithography techniques with SU-8 negative photoresist, layering each set of features with a

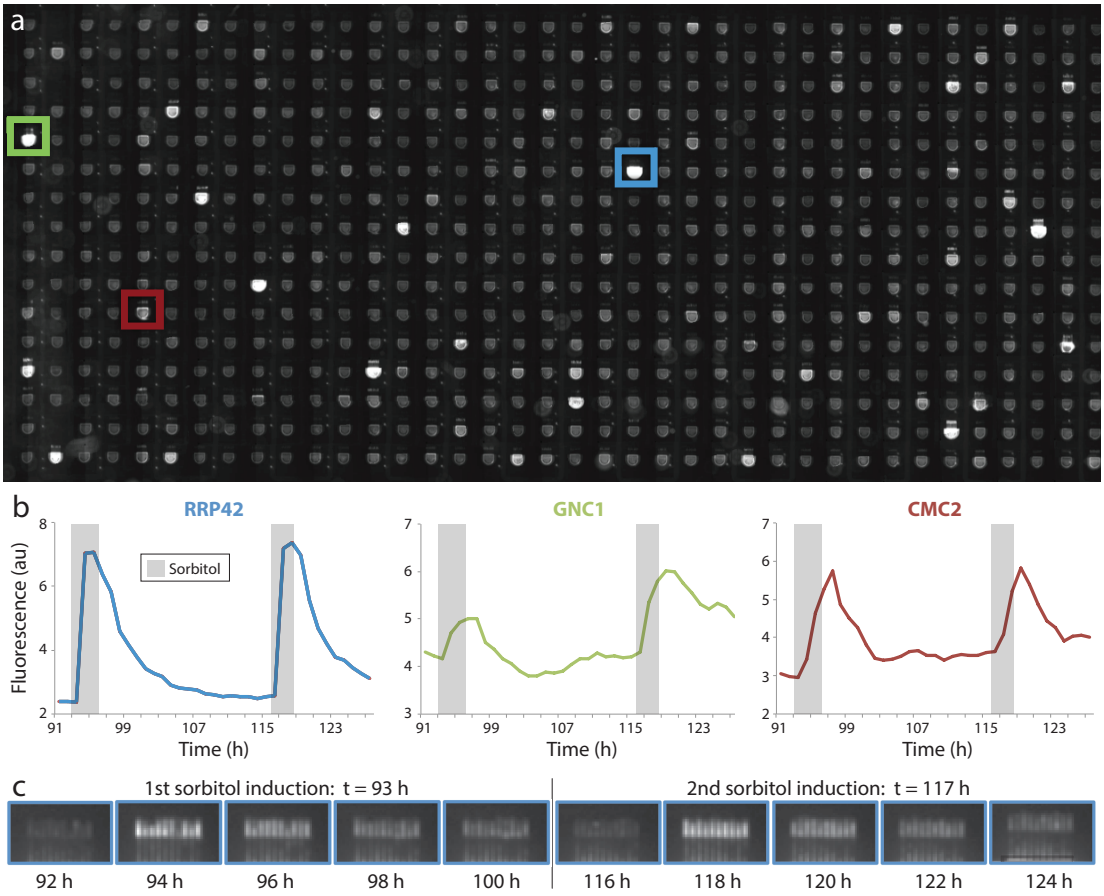

Figure 2.11: Validation of induction protocol and fluorescence measurement with *S. cerevisiae*. (a) An image of a region of the large-scale device showing approximately 608 strains just before an induction with sorbitol. (b) The temporal responses of strains in the colored squares from panel (a) show fluorescence of three strains responding to sorbitol induction (gray bars), an osmotic stress inducer. (c) Snapshots of the cell chamber region of strain RRP42-GFP taken every two hours before and after the two sorbitol inductions.

unique height. The wafer is then coated with silane to prevent SU-8 features from tearing off the wafer with successive PDMS pours.

2.5.2 PDMS device fabrication

Timing: 3 hours

1. In a clean weighing dish, mix 70 grams of the Dow Corning Sylgard 184 elastomer base with 7 grams of the Dow Corning Sylgard 184 curing agent. Mix vigorously with a clean stir rod for 5-10 minutes until the PDMS solution is well mixed.

2. Place the PDMS mixture in a vacuum desiccator to remove bubbles. Vent the vacuum desiccator as needed so that the bubbles do not spill over the weighing dish.

3. Place a 5” x 5” x 0.125” glass plate in the center of a two-layer 8” x 8” piece of aluminum foil. Carefully fold up the aluminum to create a dish around the glass plate. Overlap the aluminum foil over the edges of the glass to minimize PDMS leaking underneath the glass dish.

4. Place the patterned silicon wafer on the center of the glass plate and pour the PDMS mixture onto the center of the wafer, using a spatula if necessary, to get the viscous mixture onto the wafer.

5. Place the wafer stack into a leveled vacuum desiccator and degas until all bubbles are removed.

6. Once the bubbles are removed, use the blunt ends of two pipette tips to center the wafer if it slid to one side, and gently push down on opposite sides of the wafer to push out PDMS that seeped under the wafer.

7. Place the wafer stack into a level oven and bake at 95°C for 1 hour.

8. Remove the wafer stack from the oven. Use a razor blade to cut off the foil from the wafer stack and cut off the excess PDMS around the wafer.

9. Gently slide a razor blade horizontally between the wafer and glass plate and then remove it. Repeat this around the circumference of the wafer until the wafer separates from the glass plate. *Note*: The razor blade must slide horizontally between the glass and the wafer. Wafers are extremely fragile and if the razor blade is angled then the wafer will break.

10. Using a razor blade, remove any excess PDMS from the bottom of the wafer. Peel the PDMS off of the feature side of the wafer in the direction of the major channels.

11. Place the PDMS on a cutting mat with the feature-side up to keep the PDMS clean. Using a razor blade, cut out each PDMS device. Punch out the inlet and outlet channels using a stainless-steel puncher.

12. Rinse each device with 70% ethanol and blow it dry with compressed air or pressurized nitrogen gas.

13. Remove debris from each device using Scotch tape, cleaning the feature side four times and the non-feature side twice. Use forceps to gently press the tape into the features to remove all debris. Leave tape on each side to keep the devices clean.

2.5.3 Glass slide preparation

Timing: 1 hour

1. Sonicate glass slides in a 2% Helmanex III solution for 30 minutes at 40°C.

2. Rinse glass slides with deionized water, rubbing them with a clean latex glove.

3. Completely dry the glass slides with pressurized nitrogen gas and ensure that no streaks are visible.

4. Store glass slides in a clean, dust-free environment until used.

2.5.4 Cell preparation

Timing: 48-72 hours

1. Using Singer Plus Plates, pour agar plates with the appropriate cell culture media on a level surface, pouring 42 mL of media + agar into each plate. Allow plates to dry on the benchtop with the lids covered for 48 hours before parafilming and storing at 4°C.

2. Fill 96-well or 384-well liquid plates with overnight cultures of the strains of interest, or thaw glycerol stocks of strains of interest in a 96-well of 384-well density format.

3. Using the Singer ROTOR, spot the liquid plate onto an agar Plus Plate. Use the default pinning settings for both the source and target plates.

    - If spotting *E. coli*, grow the cells at 37°C overnight.
    - If spotting *S. cerevisiae*, grow the cells at 30°C for 2 days.

4. Using the Singer Stinger single colony arrayer, rearray the 96-agar or 384-density agar plate onto a set of 4, 1536-density plates that matches the array of the microfluidic device(s). These plates will later be combined into one, 6144-density plate that will be spotted to the microfluidic devices. *Note*: if spotting devices with fewer strains, multiple devices can be arrayed onto one set of 1536-plates.

5. Grow the cells overnight.

    - If spotting *E. coli*, grow the cells at 30°C overnight.
    - If spotting *S. cerevisiae*, grow the cells on the benchtop overnight.

    Note: 1536-density plates can be stored in the fridge and continually used as source plates for up to 6 months.

2.5.5 6144-density plate and acrylic tool preparation

Timing: 1 hour for E. coli, 25 hours for S. cerevisiae

A: If spotting *E.coli*:

1. Using the Singer ROTOR, combine the 4x 1536-density agar source plates onto one 6144-density agar plate using the "1:4 Array" program and the pinning settings listed in Table 2.2. Repeat this for two target plates.

2. Grow one of the target plates for one hour at 37°C if the source plates have been used less than three times, or for 90 minutes at 37°C if the plates have been used more than three times.

3. Using the other 6144-denstiy target plate, the "Replicate" program on the Singer ROTOR, and the pinning settings listed in Table 2.2, spot cells from 6144-density target plate onto the clean acrylic alignment tool. These cells will be used as alignment markers for the PDMS device.

B. If spotting *S. cerevisiae*:

1. Using the Singer ROTOR, replicate the 4x 1536-density agar source plates onto 4x 1536-density agar source plates using the "Replicate" program. Grow the plates at 37°C overnight.

2. Using the Singer ROTOR, combine the 4x 1536-density agar source plates onto one 6144-density agar plate using the "1:4 Array" program and the pinning settings listed in Table 2.2.

3. Using the "Replicate" program on the Singer ROTOR and the pinning setting listed in Table 2.2, spot cells from 6144-density target plate onto the clean acrylic alignment tool. These cells will be used as alignment markers for the PDMS device.

Table 2.2: Singer ROTOR pinning settings for each step of the microfluidic device loading process.

Source to target description	Pinning pressure (%)		Pinning speed (mm/sec)		Pinning overshoot (mm)	
	Source	Target	Source	Target	Source	Target
1536 agar to 6144 agar	58	64	19	10	2	1
6144 agar to acrylic	50	100	10	10	0.6	0.6
6144 agar to microfluidic device for *E. coli*	55	70	10	10	0.6	0.6
6144 agar to microfluidic device for *S. cerevisiae*	50	64	10	10	0.6	0.6

2.5.6 Aligning the PDMS to the acrylic tool

Timing: 30 minutes

A. If spotting multiple devices with fewer strains:

1. Using a photolithography mask aligner or similar system, set the acrylic tool on top of the mask holder with the alignment cells facing up. Bring the cells of one device into focus.

2. Remove the Scotch tape from one PDMS device, avoiding touching the feature side of the device.

3. Gently place the PDMS device on top of the alignment cells, feature-side up, such that the center of the spotting regions is centered over the cells.

4. Place tape on top of the PDMS, pressing the PDMS down to ensure adhesion of the PDMS onto the acrylic tool.

5. Repeat this for each device on the acrylic tool.

B. If spotting one large device with many strains:

1. Using a photolithography mask aligner or similar system, place the PDMS device feature-side-down on top of the wafer chuck. Remove the tape from the non-feature side of the PDMS.

2. Place the acrylic tool on top of the mask holder, with the alignment cells facing down, above the PDMS.

3. Using the micrometers on the mask aligner, align the center of the spotting regions with the center of the alignment cells.

4. Bring the PDMS and acrylic tool into contact using the wafer chuck.

5. Remove the tape from the feature-side of the PDMS and check the alignment, ensuring that the PDMS did not shift when it came into contact with the acrylic tool. If necessary, adjust the alignment.

6. Re-tape the feature side of the PDMS until ready to expose the PDMS to oxygen plasma.

2.5.7 Oxygen plasma exposure

Timing: 10 minutes

1. Expose the clean 4x3" glass slide and the PDMS acrylic stack to 30W of oxygen plasma for 30 seconds.

2. Blow any dust off the glass slide and PDMS acrylic stack with compressed nitrogen.

2.5.8 Cell preparation for *S. cerevisiae* spotting

Timing: 10 minutes

1. If spotting *S. cerevisiae*, using the Singer ROTOR, combine the heat shocked 4x 1536-density agar source plates onto one 6144-density agar plate using the "1:4 Array" program and the pinning settings listed in Table 2.2.

   Note: This plate should be used immediately to spot cells onto the oxygen plasma exposed PDMS device.

2.5.9 Loading and bonding the device

Timing: 10 minutes

1. Using the Singer ROTOR and the parameters listed in Table 2.2, spot the cells from the 6144-density agar plate to the oxygen plasma exposed PDMS acrylic stack.

2. Peel the spotted PDMS off the acrylic piece and gently place it face down on the center of the oxygen-plasma-exposed glass slide.

3. Gently tap the top of the PDMS, ensuring that the device bonds to the glass.

4. Incubate the device at 37°C for at least two hours.

2.5.10 Experimental set-up

Timing: 40 minutes

1. Place the bonded PMDS device in a vacuum desiccator for at least twenty minutes.

2. Prepare an inlet syringe using methods previously described by our group [16].

3. Prepare and outlet syringe (mid-scale) using methods previously described by our group [16] or prepare an outlet tube and metal connector feeding into a waste bottle.

4. Mount the bonded PDMS device on the desired imaging platform, plugging in the inlet syringe first and then the outlet tube when a bead of liquid has formed on the outlet port.

5. Image the device at the desired temporal resolution.

2.6 Acknowledgements

This chapter is currently being prepared for submission as A custom high-throughput microfluidic and optical platform for continuous culturing and measurement of *E. coli* and *S.*

cerevisiae. Csicsery, Nicholas*, Stasiowski, Elizabeth*, Thouvenin, Gregoire, Graham, Garrett, O'Laughlin, Richard, Mather, William H., Ferry, Michael, Cookson, Scott, and Hasty, Jeff. (*equal contribution) The dissertation author was one of the primary authors and researchers of this material.

Chapter 3

Dynomics: A microfluidic and machine learning platform for genome-scale transcriptional dynamics and environmental biosensing

3.1 Introduction

With a platform developed for the continuous culture and measurement of up to 2,176 microbial strains, we sought to demonstrate its utility in measuring gene expression dynamics at the genome scale. As a further application, we investigated the intrinsic ability of a fluorescent promoter library to serve as a *de facto* array of whole-cell biosensors. The development of whole-cell biosensors has been at the juncture of synthetic biology and the study of genome-scale dynamics [13, 59]. Previously, whole-cell sensors have been developed towards applications including the detection heavy metals for environmental safety [18, 59, 60], sensing metabolites relevant to bioproduction [61–64], or operating *in vivo* towards diagnostic applications [65–67].

Whole-cell sensors often harness transcriptional machinery, using either native or engineered pathways in *E. coli*, to tie specific environmental inputs to the expression of a reporter gene from a transcriptional promoter. In this chapter, I demonstrate the utility of Dynomics, a 2,176-strain microfluidic and imaging platform, for both discovery of these transcriptional elements, namely, in the case of heavy metals, where *E. coli* have already evolved native sense and response mechanisms [68, 69].

3.1.1 Heavy metal pollution and effects on human health

Heavy metal contamination in water sources is a global threat that continues to grow due to industrialization, mining, and aging water supply infrastructures [70]. Such contamination is detrimental in both drinking and agricultural water supplies, with top contaminants including arsenic, cadmium, mercury, chromium, lead, zinc, iron and copper.

Arsenic is a pervasive carcinogenic agent in groundwater globally that has been linked to skin, bladder, lung, and other cancers [71]. It's natural and widespread abundance makes it a metal of significant concern, both in the US and worldwide. Cadmium, largely introduced to the environment through industrial processes and cadmium-polluted waste, has particularly harmful effects on human health at low concentrations. Cadmium consumption leads to acute gastrointestinal effects, kidney and bone damage, and potential carcinogenic effects [72]. Mercury toxicity varies depending on organic or inorganic forms and exposure levels, but symptoms range from muscle weakness to neurological dysfunction. While mercury enters water sources through natural and human processes (coal burning and mining), its primary introduction to humans is through the consumption of fish that have biomagnified this atmospheric mercury [73]. Chromium and lead are other industrial pollutants, with chromium linked to carcinogenic effects [74] and lead, commonly entering drinking water through corroded plumbing, notable for its neurotoxicity, harming neuropsychological development in children [75, 76].

Zinc, iron, and copper are all metals which are required for many biological processes, but

can be toxic to humans in high doses. However, because human cells have developed powerful homeostasis mechanisms to these metals, acute and chronic toxicity is less common. Chronic excess zinc exposure has been shown to cause anemia and copper and ferritin deficiencies [77], excess iron can have negative effects on the liver [78] and excess copper consumption primarily affects the gastrointestinal tract and liver [79].

The contamination of US water supplies has affected hundreds of thousands of Americans in cases such as Flint, Michigan where an estimated 140,000 people were exposed to dangerously high levels of lead in their drinking water for over a year, with thousands of children testing positive for elevated blood levels for years to come [80, 81]. While Flint, MI received much of the media's attention, there are almost 3,000 small counties in the United States with a prevalence of lead poisoning twice that of Flint's [82].

Heavy metal contamination also arises from large-scale spills when dams collapse at mines, dumping toxins into nearby rivers. One recent spill was the Gold King Mine Disaster near Silverton, CO on August 5, 2015 [83]. The Gold King Mine, which was abandoned in 1923, had accumulated a large amount of contaminated groundwater inside of it, a consequence of acid mine drainage [83]. The 2015 spill occurred when the Environmental Protection Agency (EPA) went to clean the site, but accidentally caused a dam plugging the mine's entrance to fail. All of the mine's waste water, including metals such as cadmium, copper, zinc, iron, cadmium, and lead, poured into Cement Creek, and then into the Animas River, followed by the San Juan River. The thousands of members of the Navajo Nation, through which the San Juan flows, had to shut off irrigation from the river to their crops and livestock in the days after the spill, causing widespread damage and lasting effects on food supply [84].

Heavy metals have long posed a threat to both drinking and agricultural water supplies; however, sufficient equipment for measuring levels of heavy metals is not readily available. Currently, water analysis requires periodic sampling that is sent to an off-site laboratory for analysis, eliminating the ability to detect early signs of pollution or when pollution has reached

safe levels [85]. Not only is such sampling inefficient for large mine spills causing water to be shut off from agriculture for days longer than it needs to be, but such sampling occurs infrequently in municipal areas and rarely in people's homes. With the increasing concern of heavy metal pollution, a real-time, cheap, and portable solution is necessary.

3.1.2 Current methods for metal detection in water supplies

Of the most widespread current methods for heavy metal detection are atomic absorption spectroscopy (AAS), inductively coupled plasma (ICP), and anodic stripping voltammetry (ASV). AAS atomizes elements in a sample with one of several techniques (flame, electrothermal, hybrid) and measures the amount of light absorbed by the vapor solution. Inductively coupled plasma based methods use plasma to atomize a sample, then optical emission spectrometry of the excited atoms (ICP-OES) or mass spectrometry (ICP-MS). ASV is a less expensive, but lower throughput, method that coats an electrode with metals in solution, and measures current as the metals are stripped off. All of these methods are fast and incredibly precise, however, all require the collection of environmental samples that are often brought to a laboratory for analysis [86]. The lack of a reliable, real-time, and field-deployable metal sensor offer an opening for which biosensors are well suited to fill.

3.1.3 Biosensors for heavy metal detection

With the growth of synthetic biology, biosensors have become a highly researched area for detecting heavy metals. Biosensors can be broadly classified into cell-free systems and whole-cell systems. Protein based biosensors include using aptamers, nucleic acids, peptides, enzymes, antibodies, or proteins as probes for sensing heavy metals. While these systems are able to detect low concentrations of analytes quickly and specifically, they can be expensive and often require immobilization, purifications, and involved analysis [59].

The evolution of genetic engineering has made it possible for creating whole-cell biosensors that can detect samples faster, cheaper, and with less operational expertise required than traditional analytical chemistry techniques. They offer similar detection limits as cell-free biosensors but are cheaper and easier to cultivate. Genetic engineers can also harvest the systems microbes have already evolved for processing heavy metals.

Combining microfluidic technologies with whole-cell biosensors has led to a massive reduction in size and cost for in-line heavy metal detection systems. Many microfluidic biosensors to date have been single-strain, single-toxin devices or devices that require a complex system with multiple external connections, constraining these devices to a lab setting [29, 59]. Instead of single-strain devices, genome-scale microfluidic instruments that can culture, monitor, and measure thousands of strains could leverage biology's natural ability to sense and respond to environmental stimuli, resulting in a non-specific sensor that could sense any substance that perturbs a microbe. Coupling artificial intelligence (AI) algorithms with the fluorescence outputs of such a device would allow for a robust, simple, field-deployable device to sense toxins of interest in real-time that would output if that toxin is present.

3.1.4 Artificial intelligence for biological discovery

Over the last two decades, computer science has experienced a massive increase in computing and algorithmic power in the field of artificial intelligence (AI). Concurrently, the rise of -omics technologies has led to exponential increase in the volume of experimental data [87]. However, machine learning has not yet fully delivered on its potential to facilitate scientific discovery because of the "black box problem": as an algorithm's ability to model complex phenomena grows, its decision-making processes become more and more obscured from its operators. Explainable artificial intelligence (XAI) techniques have begun to address this problem, with model-specific approaches in the life sciences [88–90], and recently, a mathematically-consistent method for understanding the decision-making process of any AI classifier [91, 92]. In

the following chapter, I will describe how we use these AI advances to better leverage the data produced by Dynomics.

3.2 Deploying Dynomics

3.2.1 Dynomics platform overview

Using the device developed in Chapter 2, we have developed a high-throughput microfluidic biosensor and associated software platform, Dynomics, that can simultaneously co-culture and monitor 1,807 strains of a promoter-GFP *E. coli* library [26] under time-varying environmental conditions, successfully detecting the presence of Cu(II), Zn(II), Fe(III), Pb(II), Cd(II), and Cr(VI). Initially, time series and fold-change data are used to identify and quantify responsive strains. The data is further leveraged via both deep learning classifiers, for real-time detection of inducers, and XAI algorithms, to quantify each strain's impact on the classifier's predictions and understand which strains are responding to each metal (Figure 3.1).

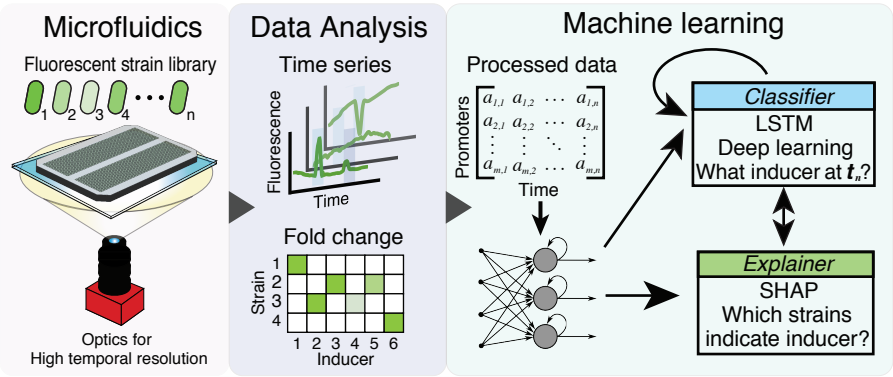

Figure 3.1: The Dynomics platform. Fluorescent strain libraries are loaded onto large-scale microfluidic devices that can be fully captured in a single image using custom optics. Parallel cultures of *E. coli* are subjected to multiple exposures of different stimuli with time series and fold-changes used to quantify responsive strains. Machine learning algorithms are trained on preprocessed data to enable real-time stimulus detection.

3.2.2 Microfluidic device optimization

Based on the initial design in section 2.2.2 the Dynomics microfluidic device was optimized for straightforward experimental setup, reliable trap filling and cell retention, and optimal fluorescence signal from each spotted microcolony (Figure 3.2a-c). The single media inlet-outlet device requires only two fluidic connections after cell spotting and chip bonding. The media inlet channel feeds a total of 2,176 4-μm-tall cell traps. Trap shape and spacing allows a 6,144 SBS-density pin pad to deposit cells into the back of the trap, where they grow towards the tapered opening interfacing with 50 μm-tall minor media-channels. These minor channels branch off of a larger 230 μm-tall major media-channel manifold system, as designed in section 2.2.2. Once spotted cells have reached confluence, inducer compounds can be pulsed in at user-specified frequencies with the dynamic response of each strain measured down to a 3-minute temporal resolution (Figure 3.2d).

3.2.3 Dynomics Optical Enclosure

The Dynomics optical enclosure is a stand-alone microscope for imaging the microfluidic device with a 70 mm x 70 mm field-of-view (Figure 3.3a). The enclosure includes an SBIG STX-16803 CCD camera (a1), a custom lens stack assembly (a2), two blue excitation LEDs (a3) and associated drivers (a7) for GFP imaging, a green LED (a5), associated driver (a7), and diffuser stack assembly (a4) for transmitted light imaging, a temperature-controlled enclosure where the microfluidic device is mounted (a4), all necessary power supplies and wiring (a8), and a Tegra computer with custom software (a7) to control the LEDs, imaging, temperature, and to sync the images onto servers via WiFi. The enclosure costs about $15,000, compared to an off-the-shelf microscope that can cost an order of magnitude more. The optical enclosure images in both transmitted light and fluorescence channels with a 36 μm optical resolution, comparable to a lab-grade microscope at 4x magnification (Fig. 3.3c-e). With a nominal imaging frequency

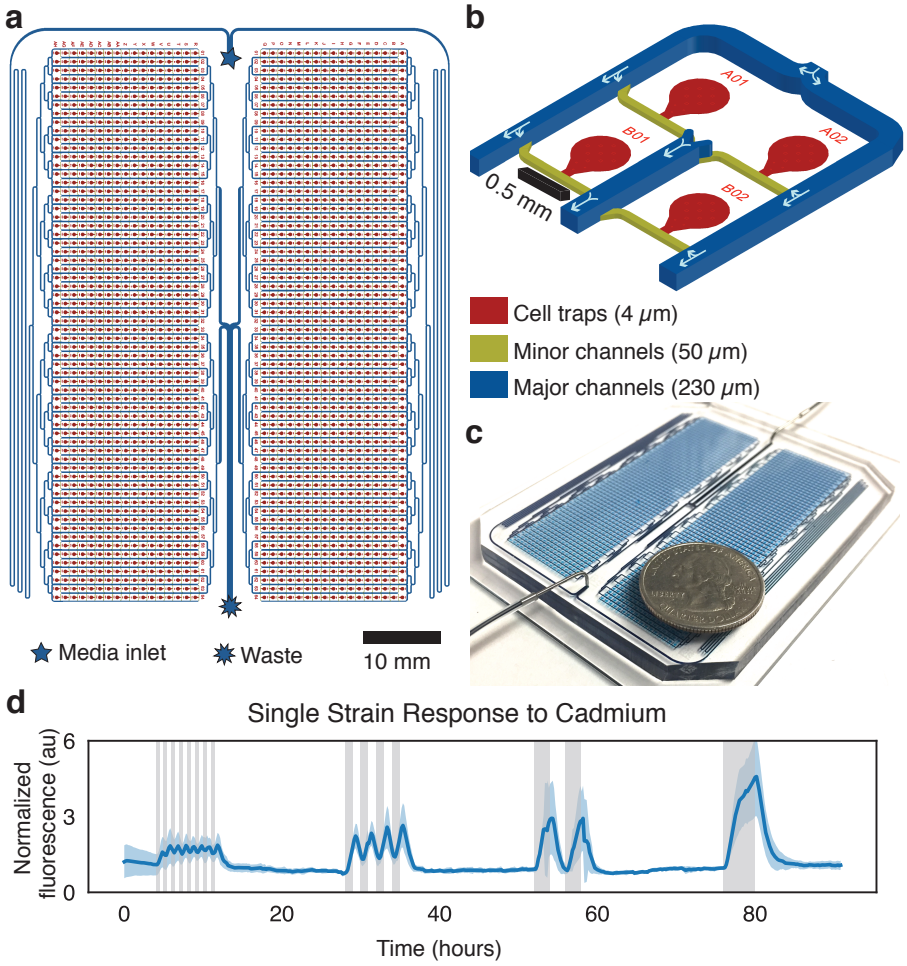

Figure 3.2: The Dynomics microfluidic device. a) Design of the Dynomics 2,176-strain microfluidic device with cell traps in red and media channels in blue and yellow. b) Detailed schematic of four strain banks with arrows showing direction of media flow. c) Mean fluorescence (solid blue) and standard deviation (shaded blue) of the E. coli *zntA* promoter driving GFP to repeated cadmium inductions (gray bars) with periods increasing from left to right (30 min., 2 hrs., 4 hrs., and 8 hrs.)

of 3 minutes, Dynomics allows us to adjust environmental conditions and measure the resulting changes in gene expression at a high temporal resolution.

3.2.4 Screening for responsive promoters to heavy metals

Using the Dynomics platform with a previously developed GFP *E. coli* promoter library [27], 1,807 unique *E. coli* promoters were screened against nine heavy metals (Cu(II), Zn(II),

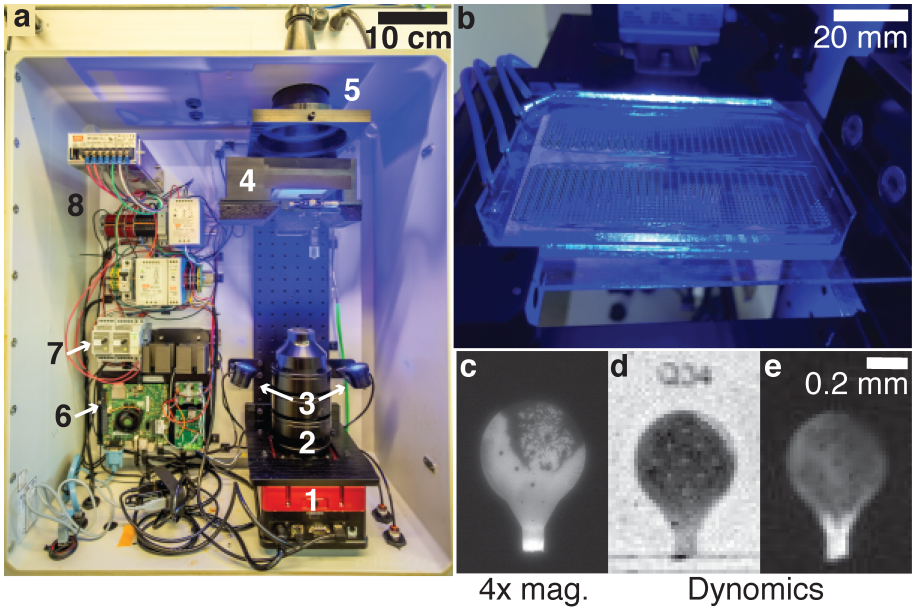

Figure 3.3: The Dynomics custom optical enclosure. a) Components include (1) a SBIG STX-16803 CCD camera, (2) a custom lens stack assembly, (3) blue excitation LEDs for GFP fluorescence imaging, (4) a temperature-controlled enclosure where the microfluidic device is mounted, (5) a green LED and associated diffusers and lenses for transmitted light imaging, (6) a Tegra computer and software for controlling imaging and transferring images to a server, (7) LED drivers, and (8) all associated power supplies and wiring. b) An image of a large-scale microfluidic device mounted in (a4). c-e) The Dynomics microfluidic device imaged on a standard research grade scope with 4x magnification, and on the Dynomics enclosure for both transmitted light (d) and fluorescence (c,e) channels. The Dynomics enclosure has 36 μm resolution.

Fe(III), Pb(II), Cd(II), Cr(VI), Hg(II), As(III), Sb(III)) at environmentally relevant concentrations (Table 3.4). Screening experiments lasted 7-14 days, with cells exposed to a different heavy metal every 24 hours (Figure 3.4).

Promoters responsive to each metal can be identified through a combination of clustering and fold-change analysis. A high-level view of the 1,807 promoter time traces (Figure 3.5a) and subsequent clustering (Figure 3.5b) reveal distinct classes of transcriptional responses to a single four hour zinc exposure. In (Figure 3.5b), clusters 1 and 2 include promoters that are up- and down-regulated, respectively, in the presence of zinc, but return to baseline expression levels within 15 hours of zinc removal. Clusters 3 and 4 include promoters that are up- and down-regulated, respectively, with slower dynamics. Gene ontology (GO) enrichment analysis

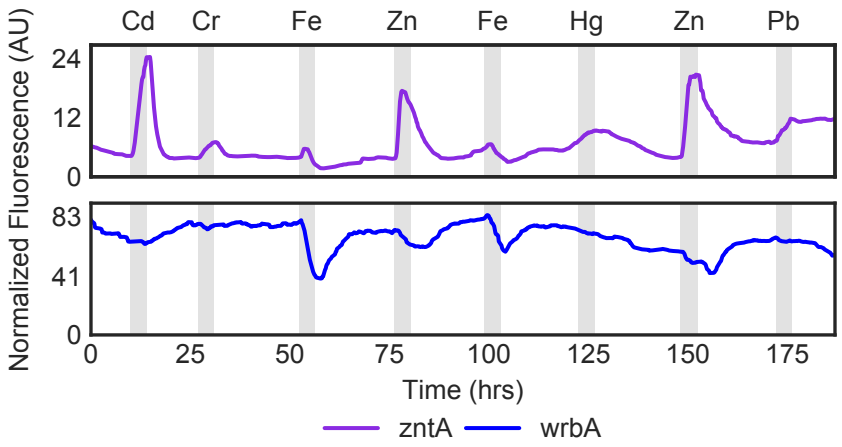

Figure 3.4: Responsive strains over the duration of a Dynomics experiment. Normalized fluorescence for two strains is plotted over the duration of one experiment, with four-hour heavy metal inductions (gray bars) occurring once daily.

suggests that from these four clusters, genes associated with cellular stress are up-regulated (cellular detoxification and antibiotic metabolic process) while genes involved in metabolism and biosynthesis are down-regulated (nitrogen metabolism and glutamine family amino acid biosynthesis). The full list of GO terms and their associated p-values are listed in Table 3.1.

Individual responsive strains for each metal were identified, based on their fold-change response (Figure 3.6a) to four-hour metal exposures, which were repeated in random order once every 24 hours. Fold-change detection highlights the promoters displaying the strongest response to each metal. Subsequent investigation of the most responsive strains (Figure 3.6b) quantitatively elucidates dynamics, such as amplitude, relaxation time, and response rate, all of which important factors for their use in the study of gene expression regulation and continuous biosensing applications. While many of the Dynomics-identified sensing strains such as *zntA* [93] or *cueO* [94] have well documented metal interactions, others are less studied or poorly annotated, particularly members of *E. coli* 'y-ome' [95]. Overall, the methods of data analysis reveal each metal to have responsive promoters with a unique signature of up- and down-regulation. However, in view of the ultimate task of a biosensor, which is to distinguish the presence of a metal based

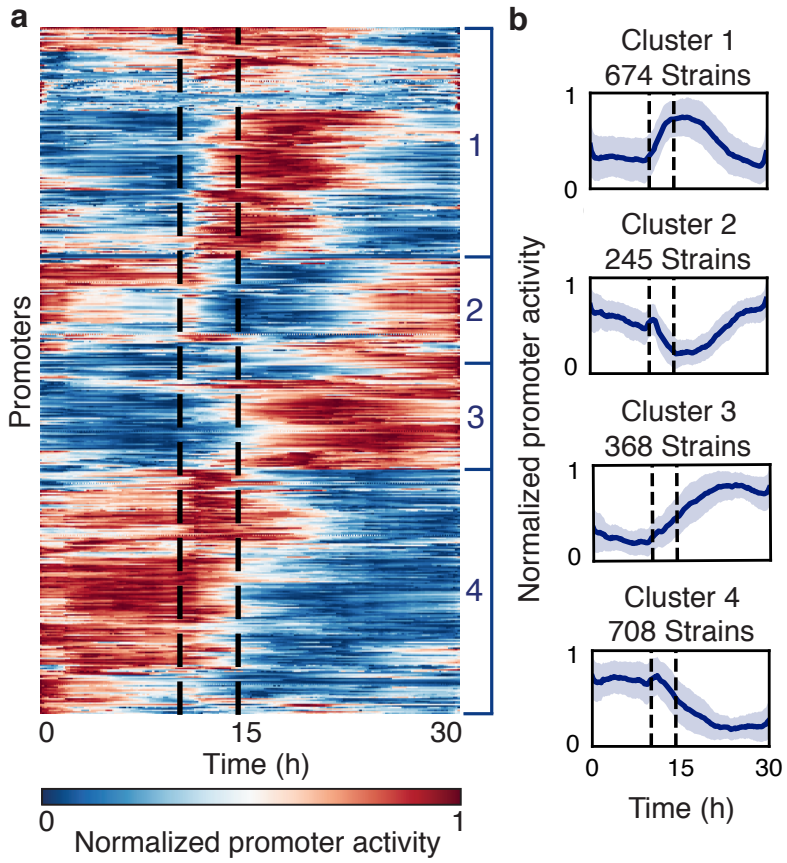

Figure 3.5: Dynomics as a screening tool for heavy metal responsive promoters in *E. coli*. a) Fluorescence response of an *E. coli* promoter library during a four hour 50 ppb Zn induction (dotted window). Each row represents the promoter activity, normalized between 0 and 1, of a single strain, with 1,995 total strains represented. Four clusters from agglomerative clustering are labeled on the right. b) Four clusters of strains calculated from agglomerative clustering from the data in panel a. The mean (dark blue line) and standard deviation (dark blue shading) of all strains in each cluster is plotted. The dotted window denotes when zinc was present.

on a real-time transcriptional data, fold-change alone is difficult due to promoter non-specificity, cross-talk, noise, and low amplitude responses.

Table 3.1: Gene ontology enrichment analysis of the clusters obtained by agglomerative clustering of Figure 3.5.

Cluster	GO biological process complete	Fold Enrichment	P-value
1	localization (GO:0051179)	0.74	4.11E-02
	transport (GO:0006810)	0.71	3.30E-02
	transmembrane transport (GO:0055085)	0.62	4.47E-02
2	protein autoprocessing (GO:0016540)	9.7	4.59E-02
	transcription antitermination (GO:0031564)	9.7	4.59E-02
	carbon utilization (GO:0015976)	9.7	4.59E-02
	cellular response to light stimulus (GO:0071482)	9.7	4.59E-02
	cellular response to radiation (GO:0071478)	9.7	4.59E-02
	regulation of single-species biofilm formation (GO:1900190)	9.7	4.59E-02
	cellular response to UV (GO:0034644)	9.7	4.59E-02
	bacterial transcription (GO:0001121)	9.7	1.29E-02
	one-carbon metabolic process (GO:0006730)	7.28	2.11E-02
	tetrahydrofolate metabolic process (GO:0046653)	5.54	1.43E-02
	folic acid-containing compound metabolic process (GO:0006760)	4.31	2.68E-02
	pteridine-containing compound metabolic process (GO:0042558)	3.88	3.48E-02
	positive regulation of nitrogen compound metabolic process (GO:0051173)	2.11	3.44E-02
	macromolecule metabolic process (GO:0043170)	0.66	4.19E-02
	transmembrane transport (GO:0055085)	0.33	4.02E-02
3	thiamine-containing compound biosynthetic process (GO:0042724)	5.29	2.57E-02
	cellular detoxification (GO:1990748)	5.29	2.57E-02
	thiamine biosynthetic process (GO:0009228)	5.29	2.57E-02
	thiamine-containing compound metabolic process (GO:0042723)	4.23	4.07E-02
	thiamine metabolic process (GO:0006772)	4.23	4.07E-02
	antibiotic catabolic process (GO:0017001)	3.53	2.15E-02
	detoxification (GO:0098754)	3.53	2.15E-02
	cellular response to toxic substance (GO:0097237)	3.53	2.15E-02
	reactive oxygen species metabolic process (GO:0072593)	3.31	4.21E-02
	tetrapyrrole biosynthetic process (GO:0033014)	3.17	2.99E-02
	tetrapyrrole metabolic process (GO:0033013)	3.17	2.99E-02
	response to toxic substance (GO:0009636)	2.91	5.74E-03
	antibiotic metabolic process (GO:0016999)	2.8	1.46E-02
	vitamin biosynthetic process (GO:0009110)	2.65	1.29E-02
	water-soluble vitamin biosynthetic process (GO:0042364)	2.65	1.29E-02
	water-soluble vitamin metabolic process (GO:0006767)	2.43	1.78E-02
	vitamin metabolic process (GO:0006766)	2.43	1.78E-02
	drug catabolic process (GO:0042737)	2.41	2.54E-02
	drug metabolic process (GO:0017144)	2.17	1.33E-03
	positive regulation of biological process (GO:0048518)	0.38	4.72E-02
4	arginine biosynthetic process (GO:0006526)	2.79	4.34E-02
	arginine metabolic process (GO:0006525)	2.54	4.22E-02
	glutamine family amino acid biosynthetic process (GO:0009084)	2.36	3.63E-02
	drug transport (GO:0015893)	1.77	4.98E-02
	transmembrane transport (GO:0055085)	1.48	2.65E-02
	localization (GO:0051179)	1.28	4.48E-02
	drug metabolic process (GO:0017144)	0.5	2.90E-02

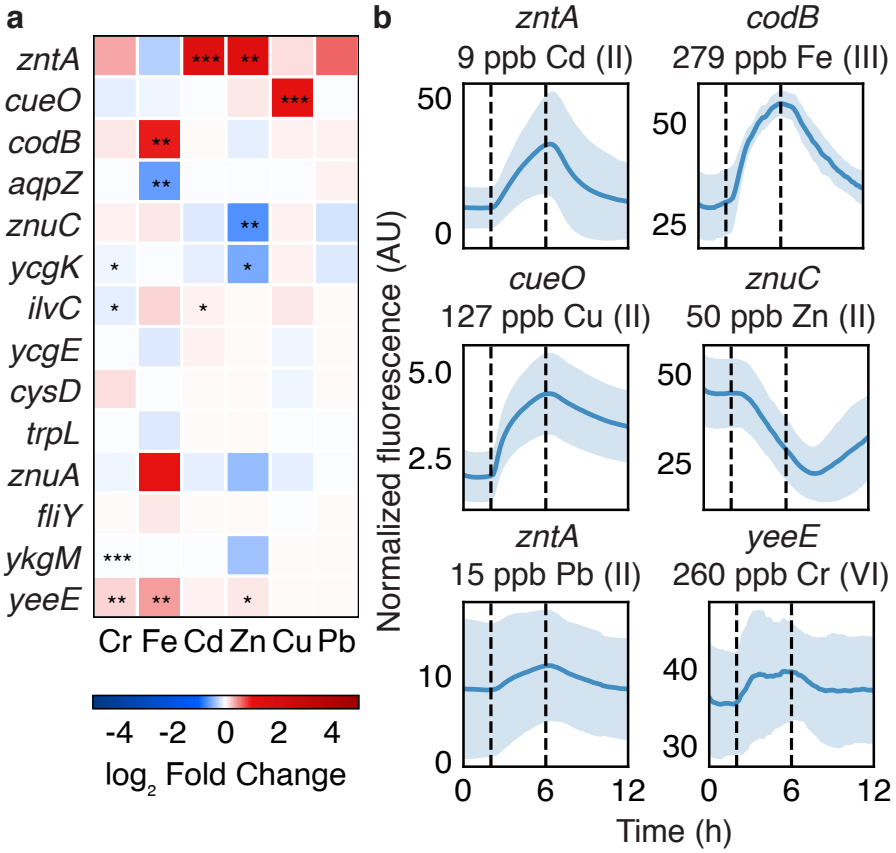

Figure 3.6: a) Fold change for top responding strains to all metals. Log2 of the average fold change is shown for the top responding strains to each heavy metal. *, **, *** indicate p-values of 0.05, 0.01, 0.001, respectively. b) Significant single strain normalized fluorescence response (blue line) and standard deviation (blue shading) across all inductions for a given metal (dotted window).

3.3 Using machine learning to leverage Dynomics data

3.3.1 Supervised machine learning

To better discriminate between *E. coli's* response to the heavy metals used in our screening, we trained and tested two types of machine learning models on the Dynomics data. The first model, known as extreme gradient boosted trees (XGBoost), is a popular decision tree ensemble-based classifier known for its ability to learn nonlinear models [96]. The second, known as a long short-term memory recurrent neural network (LSTM-RNN), is a DNN [97] selected because of

its ability to effectively utilize sample sequence history to classify time series data, a property not shared by XGBoost.

Both classification algorithms outperformed random guessing of the majority class (no toxin) on the standardized experiments' feature set, with the LSTM-RNN performing the best overall (Figures 3.7, 3.8). The LSTM-RNN was able distinguish both biotic and xenobiotic metal-spiked water from pure water with a high level of reliability. We believe this is the first instance of a multi-class classifier successfully leveraging genome-wide transcriptional dynamics to predict exposure of a biological organism to an environmental stressor.

The LSTM-RNN found iron and copper to be easily-detectable biotic metals, which is not surprising given their importance to *E. coli* cellular function [94, 98]. Cadmium was the most readily detected xenobiotic metal with the LSTM-RNN classifier, though it was sometimes confused with zinc. *E. coli* are known to use the same sensing and transport systems to capture and export excess amounts of these two metals, which possess the same number of valence electrons [93, 99]. Most classification errors occurred at the beginning and the end of the experiments' induction periods, when the LSTM-RNN occasionally had difficulty determining the exact time that each metal was added or removed from the media (Figure 3.9).

3.3.2 Explainable artificial intelligence

Taking advantage of recent XAI advances, we trained a Shapley additive explanations (SHAP) XAI on both our XGBoost and LSTM learners [91, 100]. Viewing both SHAP values (impact on classifier output) and feature values (data fed to the classifier) with respect to time offers insight into how the classifier operates in real-time (Figure 3.10). The cause of misclassification is made clearer, as SHAP dynamics reveal that the predictive impact of a strain often varies within an induction window, particularly at its start and end. Furthermore, we see how some promoters, such as *zntA*, positively contribute to the detection of multiple metals, which causes the classifier to rely on promoters with less-pronounced responses to distinguish the exposed

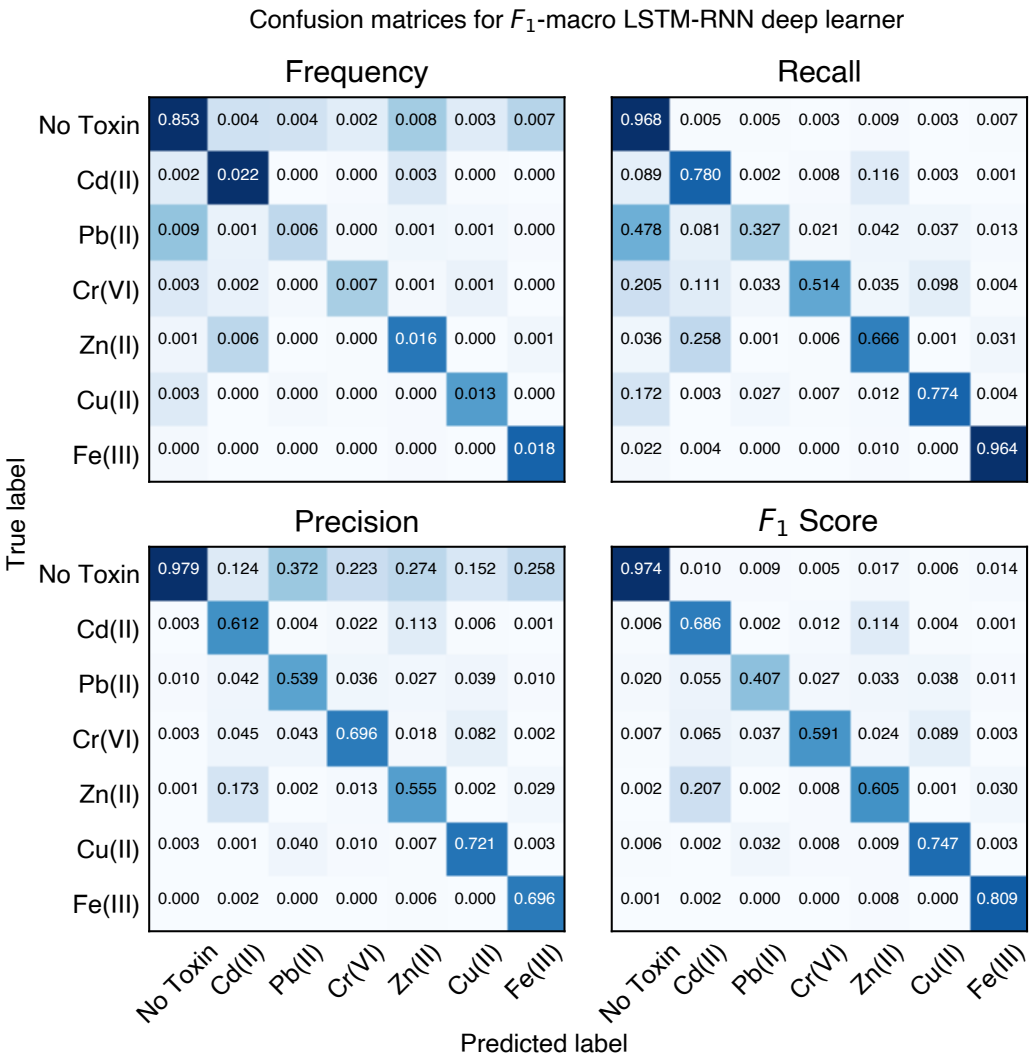

Figure 3.7: Confusion matrices showing the frequency, recall, precision, and F1 score of the LSTM-RNN classifier in predicting six metals across all experimental data (14,332 time points).

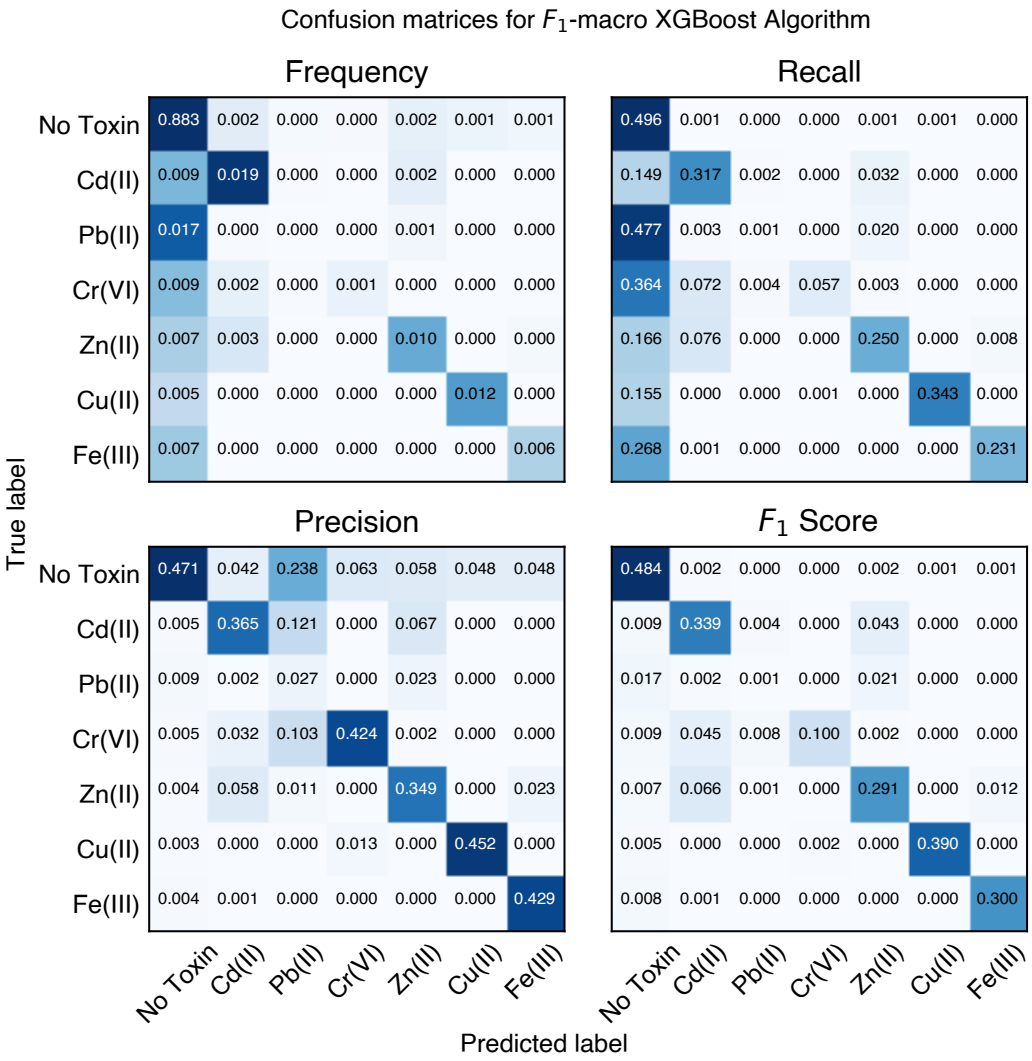

Figure 3.8: Confusion matrices showing the frequency, recall, precision, and F1 score of the XGBoost classifier in predicting six metals across all experimental data (14,332 time points).

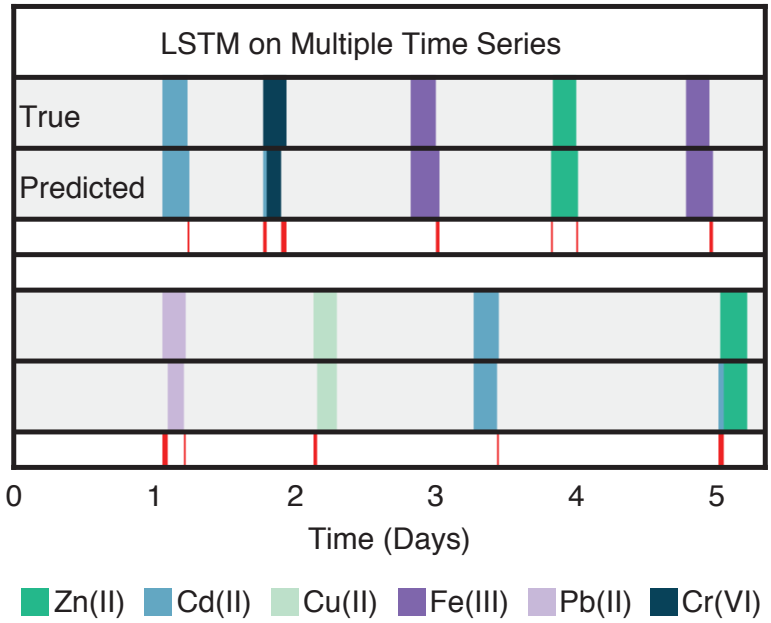

Figure 3.9: LSTM-RNN classifier applied to time series data for all six detectable metals in two different experiments. Misclassified time points are shown in red.

metal. As a result of the combination of experimental noise and weaker responses, the SHAP values of these differentiating strains are often smaller and vary more with time, explaining the misclassification of the metals. Finally, promoters that may not have been identified as responsive using fold-change analysis because of subtle, low-amplitude, and noisy responses can be identified via XAI. While these subtle responders may not serve as stand-alone biosensor strains, they provide promising targets for future sensor engineering efforts. These insights highlight the ability of the LSTM-RNN classifier to compile the influence of many strains, prominent and subtle, to make a more often than not accurate prediction of the present metal exposure.

The SHAP-XAI highlights the similarities and differences between how the LSTM-RNN and XGBoost make decisions. Figure 3.11 shows the 15 promoters with the highest mean impact on the model and the promoterless strain U139. Both methods rely heavily on the metal-sensing promoter *zntA* for the detection and discrimination of multiple metals, especially cadmium and zinc. Beyond *zntA*, XGBoost relies heavily on single strains to detect single metals, in a manner

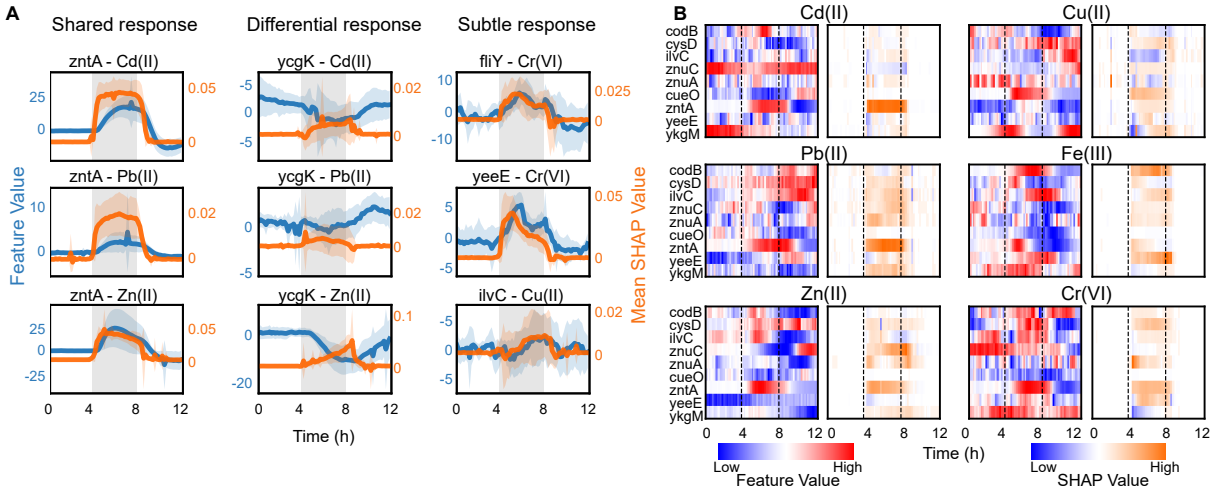

Figure 3.10: Dynamic SHAP and feature values during metal exposures. (A) Feature (blue) and SHAP (orange) time trajectories for individual promoters during metal exposures. Solid lines show the mean value over all inductions for that metal and shaded regions around lines represents standard deviation. Shaded regions represent the metal exposure window. While some promoters are responsive to many different metals, additional information from other promoters helps the classifier to differentiate each metal. Many promoters with noisy and subtle metal responses also contribute to classifier performance. (B) Feature (red and blue) and SHAP (orange and blue) values for 9 of the most information rich promoters, as determined by XAI, during metal exposures.

comparable to human attention patterns. While XGBoost is not as proficient as the LSTM-RNN classifier at predicting metals, coupled with SHAP-XAI, XGBoost is able to identify top responding strains for each metal. The LSTM-RNN, on the other hand, utilizes many strains of moderate influence in a combinatorial fashion; this tendency to find a different representation than that of the human visual system has been noted in other works [101].

When looking by class, we see overlap with some, and more distinction with others, like copper (Figure 3.12).

The ability of the XAI-coupled classifiers to identify promoters involved in metal response serves as a valuable scientific tool, suggesting potential pathways for further investigation. This is highlighted by looking at a subset of the ten most-impactful promoters individually for cadmium and iron inductions (Figure 3.13). These summary plots illustrate how the two classifiers make similar decisions through different methods. In the case of cadmium, *zntA* plays a significant

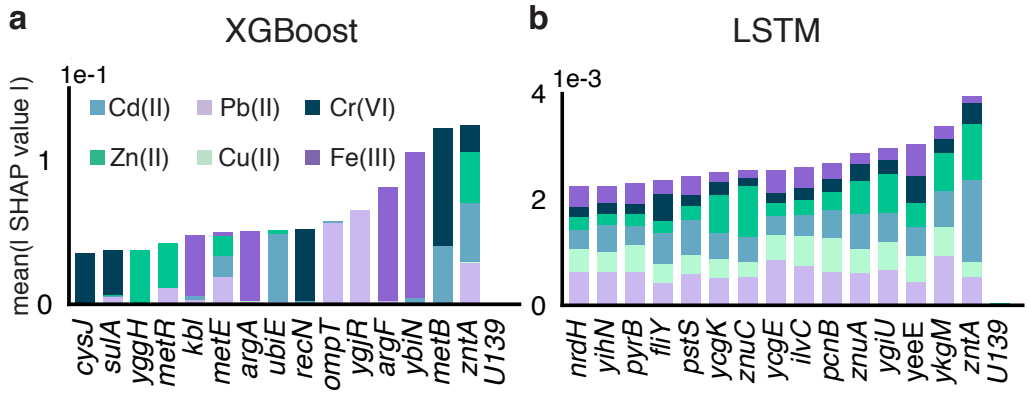

Figure 3.11: Explainable machine learning reveals *E. coli* transcriptional elements contributing to metal classification. The bar plots show the cumulative contribution based on the Shapley Additive Explanations (SHAP) values of 15 top promoters and a negative control (promoterless strain U139) to the prediction of each metal for both XGBoost (a) and LSTM-RNN (b) classifiers. Colored bars for each metal represent the mean absolute SHAP value over all experimental time points.

role for both classifiers, while different sets of genes involved in ion transport or amino acid synthesis are identified for each. Most notably, the *metE* and *metB* promoters which are involved in methionine synthesis, an amino acid known to chelate cadmium, are identified by XGBoost, while the LSTM-RNN only uses the *metE* regulator, *metR*, for detection. Similarly, with iron, we see XGBoost rely on members of the arginine synthesis, *argA* and *argC*, while the LSTM-RNN relies on different promoters that are involved in other metabolic or biosynthetic processes.

3.3.3 Biosensor validation

Given the severe impact of heavy metals on human health [102] and the persistence of water quality issues in the US [103] we sought to deploy the Dynomics platform as a real-time water quality biosensor. To validate this device for heavy metal sensing outside of a laboratory setting, we conducted experiments with media made from municipal water samples from San Diego, Seattle, Chicago, Miami-Dade, and New York City with added cadmium. Figure 3.14a shows the LSTM-RNN classifier predictions for cadmium exposures on each city's water supply. While there is some misclassification of cadmium for zinc, there are few instances of incorrectly

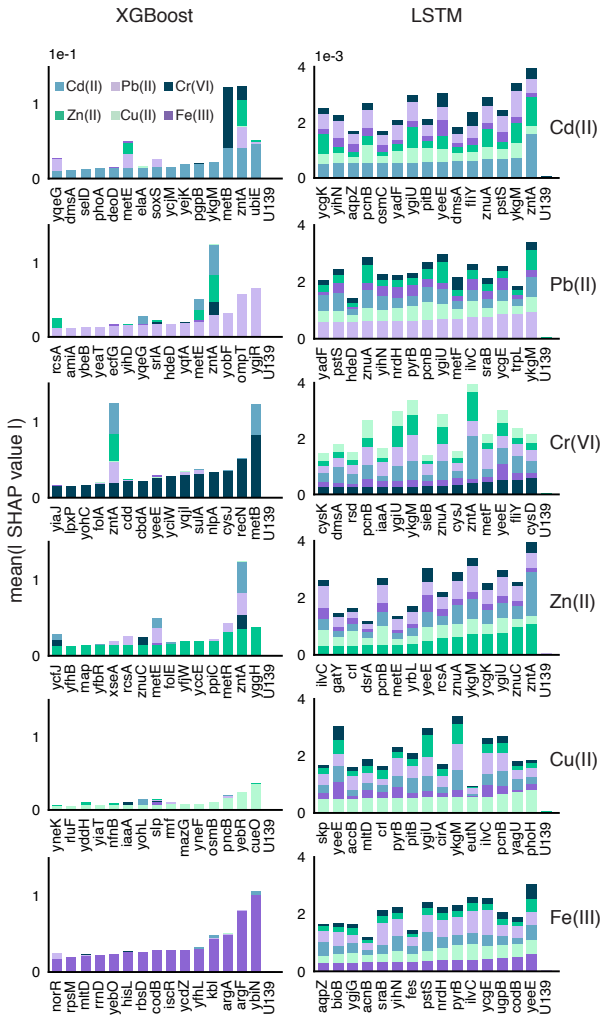

Figure 3.12: *E. coli* transcriptional elements with the highest contribution to the prediction of each metal. The bar plots show the cumulative contribution based on the Shapley Additive Explanations (SHAP) values of 15 top promoters and a negative control (promoterless strain U139) to the prediction of each metal for both XGBoost (a) and LSTM-RNN (b) classifiers. Colored bars for each metal represent the mean absolute SHAP value over all experimental time points.

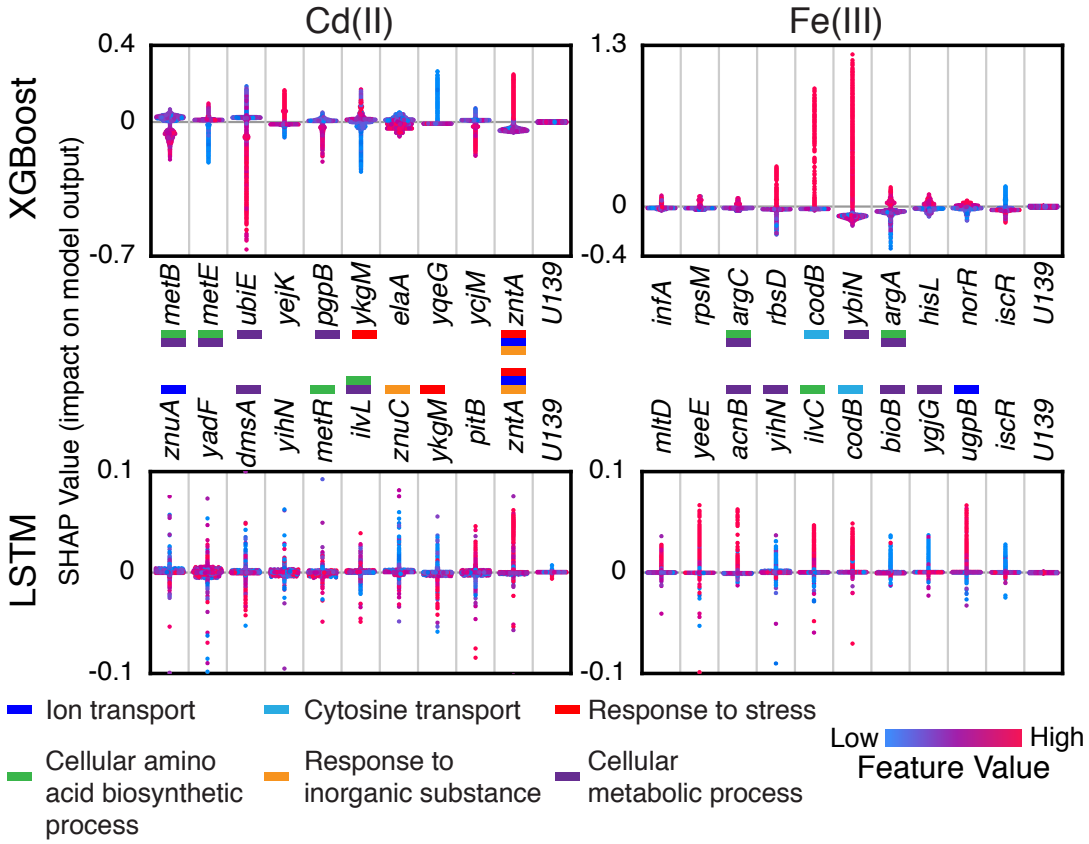

Figure 3.13: SHAP values for the top 10 promoters for Cd(II) and Fe(III) with enriched Gene Ontology (GO) terms. SHAP values are shown for 10 top promoters and a negative control (promoterless strain U139) for Cd(II) and Fe(III) for XGBoost and LSTM-RNN. Each point represents the feature value (normalized first derivative) at a time point. Positive SHAP values suggest that a given metal is present while negative values suggest its absence. Upregulated promoters (*zntA*, *codB*) give high SHAP values when feature values are high. Promoters are annotated with prominent gene ontology terms enriched between the two data sets.

predicting the presence of a toxin versus water, even with largely different water compositions between each city.

The Dynomics device was also exposed to samples collected from the Gold King Mine Spill in August 2015. Figure 3.14b shows the predictions of the LSTM-RNN classifier on samples from the spill, collected from the San Juan River. The classifier predictions are output as multi-class, multi-label probability vectors. As the sample was introduced onto the device, the probability of water decreased significantly while the probabilities of the other metals increased.

The metal with the highest probability, iron, was also the most abundant metal in the samples as measured by ICP-MS (Table 3.2). Despite the classifier not being trained on combinations of metals, nor at the concentrations present in these samples, the ability to reliably report the presence of the most prominent metal, and, to a lesser degree, the less abundant metals, suggests the broad applicability of this platform for heavy metal detection.

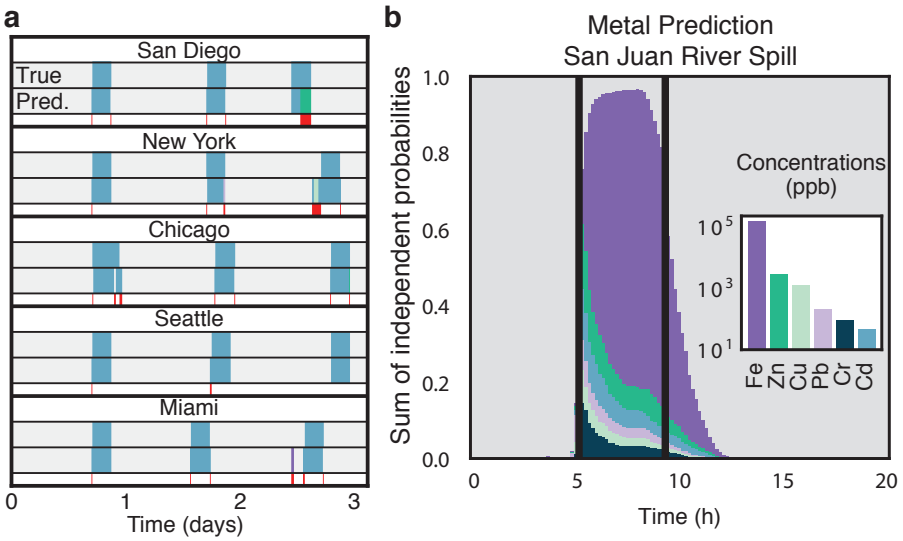

Figure 3.14: Dynomics and machine learning on environmental samples. a) LSTM-RNN classification of cadmium contamination added to five different urban water sources. The colors correspond to the metals in the inset in (b). b) Multi-class, multi-label classification of water samples from the San Juan River during the 2015 Gold King Mine waste water spill. Independent probabilities of each class are determined by the sigmoid activation function. The plot shows the sum of the classifier probabilities, averaged across triplicate sample exposures (addition and removal at vertical black lines). Inset bar chart shows the concentration of detectable metals in San Juan River samples as determined by ICP-MS. The colors of predicted toxins correspond to the metals plotted in the inset.

3.3.4 Dimensionality reduction for data visualization

To better visualize the data from 2,000+ strain Dynomics experiments, data reduction techniques were applied to the data set. Principal component analysis (PCA) was performed on induction data across several experiments to reduce the 2,176 strain dimensions into a 2D or

Table 3.2: Concentration of metals in HM9 media made with San Juan River samples as measured by ICP-MS at the Environmental Complex Analysis Lab at UC San Diego.

Metal	Concentration
Aluminum	100.9 ppm
Vanadium	1.263 ppm
Chromium	88.83 ppb
Manganese	828.3 ppm
Iron	151.8 ppm
Cobalt	1.289 ppm
Nickel	1.634 ppm
Copper	1.215 ppm
Zinc	2.947 ppm
Gallium	63.73 ppb
Arsenic	64.94 ppb
Silver	22.45 ppb
Cadmium	48.16 ppb
Thallium	33.52 ppb
Lead	209.3 ppb

3D space. We found that this data reduction technique can be used to visualize differences in the device state between water and toxin inductions. Principal component analysis (PCA) was performed on the collected data using the scikit learn library in Python. To perform PCA, the collected time series are processed using a median high-pass filter and individual time points are selected from the start and end of each induction window. Time points collected outside of the induction window are labeled as water, while those collected within the induction window are labeled as the inducing toxin. A PCA model is fit to an m-by-n matrix containing the collected data, where m is the number of time points and n is the number of strains. The model is then used to apply an orthogonal transformation to a data-set sharing the same number of strains, but not necessarily the same number of time points. The resulting matrix will maintain the number of time points, while the strain dimension will have been reduced. The reduced dimensions emphasize variance in the original data set, with the first principal component containing the highest variance.

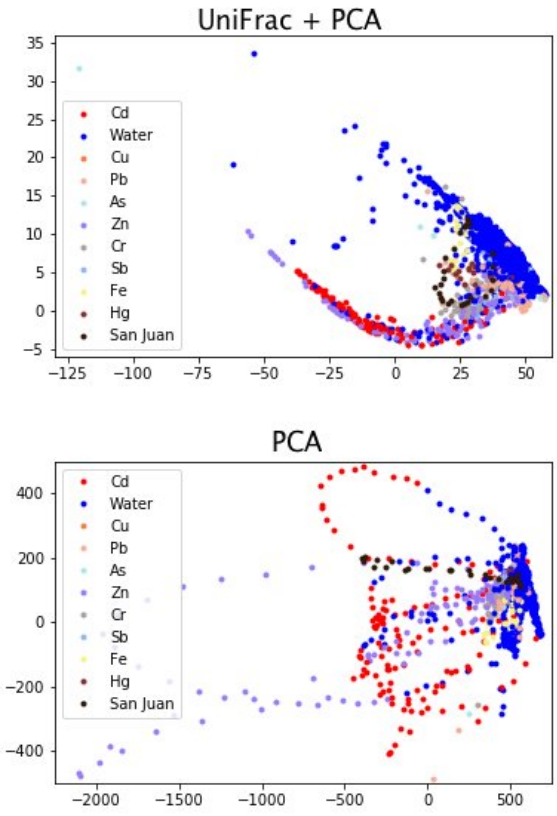

Figure 3.15: Dimensionality reduction of Dynomics data. Principal component analysis was performed on regularly processed data (bottom) and data processed with the UniFrac metric (top). Each point represents an experimental time point, with color indicating the true label. The water label indicates media with no added metal or toxin.

To improve visualizations from PCA, we adopted a method from microbiome research, UniFrac, to operate with our data [104]. UniFrac is a distance metric which accounts for the phylogenetic distance between species in metagenomic data sets. Processing data with UniFrac before dimensionality reduction emphasizes trends at the taxon level rather than the species level. To apply to our data, a gene ontology tree (analogous to a phylogenetic tree) was constructed based on the shared gene ontology labels of each gene in the *E. coli* genome. UniFrac then computed the distance between each time point with weight factored in for ontological difference between each strain. Applying the UniFrac distance metric before reducing the dimensionality of our data to 2D space resulted in much cleaner clustering of the results, and better overall

visualization (Figure 3.15). However, when simultaneously looking at all experimental data, the metals that incite the highest magnitude response (Cd and Zn) tend to be those that contribute the most to variance in the data, and this are highlighted in PCA.

Ultimately, we found that PCA produced the best results when applied to a reduced set of responsive library device positions and a reduced set of compounds. Using information from the XAI, we selected the top 10 responsive strains for each of the three metals: cadmium, copper, and iron (Figure 3.16). With overlapping response strains for each metal, a total of 28 strains were reduced to 5 dimensions, with principal components 2, 3, and 5 expressing specific variability for one of the three metals.

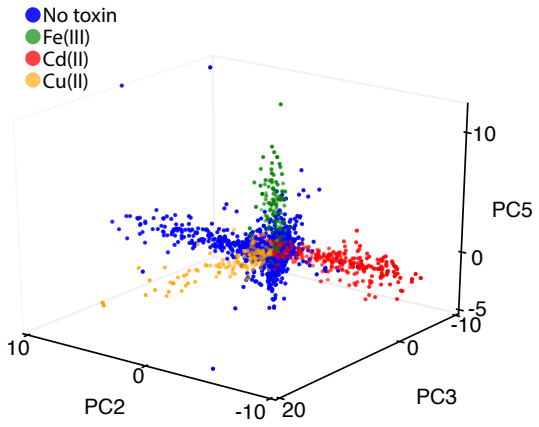

Figure 3.16: Dimensionality reduction on a reduced strain set. PCA was performed on first-derivative like features for the top 10 responsive genes, as determined by XAI, for cadmium, copper, and iron. Principal components 2, 3, and 5 are shown.

While PCA and other dimensionality reduction techniques do not replace supervised machine learning for predicting the chemical state of water, it can be a useful tool for data visualization. A particularly interesting challenge with this data set is the evolution of feature values over time. We can see from these visualization methods that the state of the device has a transition region from toxin to no-toxin, rather than an immediate jump to a segregated region of the principal component space. This further highlights challenges that the machine learning algorithms faced while classifying metals: metals are difficult to predict upon immediate

introduction. Nonetheless, after extended time on a specific metal, the expression state of the device changes in a way such that we can distinguish the presence of these metals.

3.4 Discussion

Adaptation to rapidly changing environments and external stressors is a hallmark of bacterial life and requires critical dynamic properties, many of which can be traced down to rapid and specific control of gene expression [105]. Transcriptomic technologies [106–108] have identified genes involved in environmental stress response and, more recently, furthered understanding of their mode of regulation. In this work we developed a high-throughput microfluidic platform to track the transcriptional dynamics of thousands of *E. coli* genes in parallel. Our Dynomics platform provides a temporal resolution and degree of experimental control that alternative experimental techniques do not possess [29]. In a high-throughput screen using Dynomics, we simultaneously exposed the 1,807 strains of the promoter-based *E. coli* GFP library to nine different heavy metal stressors. The fine-grained temporal gene-expression data it produced highlighted the unique dynamics of stimuli-specific genes previously identified as heavy-metal responsive [107], and allowed for the identification of gene clusters based on temporal response. Our platform possesses the genome-scale coverage and high sampling frequency needed to probe bacterial gene regulatory networks and screen large strain libraries for motifs difficult to discern using static gene expression data, yet ubiquitous in biology such as nonlinear interaction patterns and feedback loops [42].

We further illustrate our platform's capacity for exploring the dynamics of transcriptional networks by applying machine learning techniques to detect stimuli-specific features of heavy metal stress response. Here we demonstrate that supervised machine learning can leverage the dynamic change in gene expression to infer the presence of environmental stressors in real-time. Algorithms derived from two formally distinct machine learning frameworks were trained on the

time traces of the 1,807 promoter strains and successfully differentiated between metals belonging to biotic (Cu(II), Zn(II), Fe(III)) and xenobiotic classes (Pb(II), Cd(II), Cr(VI)). Deep-learning based methods, known for their ability to leverage time-series data, were more successful. In addition, we exploit recent progress in XAI to gain insight into the biological mechanisms that characterize transcriptomic adaptation to stress.

Finally, we demonstrate the real-world applicability of our platform by testing it on cadmium-spiked water from five municipal areas around the US with varying water composition, as well as water from the Colorado Gold King Mine Spill, possessing multiple metals at concentrations vastly different from those in the training sets. In both cases the machine learning algorithms successfully predicted the presence of the appropriate heavy metals. Approaches such as these involving the combination of high throughput microfluidics and machine learning may outperform single purpose biosensors in accuracy and robustness, and be adaptable to more varied sensing tasks.

3.5 Materials and methods

3.5.1 Wafer Fabrication

The silicon wafer was fabricated using standard photolithography techniques previously described by our group [16]. The cell trap layer was fabricated using 2005 SU-8 photoresist with a spin speed of 5000 rpm and had a resulting height of 3.85-4.05 μms, the minor channel layer was fabricated using 2075 SU-8 photoresist with a spin speed of 4250 rpm and had a resulting height of 45-50 μms, and the major channel layer was fabricated using 2075 SU-8 photoresist with a spin speed of 1200 on top of an undeveloped major channel layer, resulting in a final height of 200-260 μms.

3.5.2 Microfluidic device fabrication

The microfluidic devices were fabricated, loaded, and bonded with the *E. coli* promoter library [27] using the protocol in section 2.5.

3.5.3 Experimental protocol

Microfluidic experiments were performed on the Dynomics custom optical enclosure. Continuous imaging occurred every ten minutes, imaging both the transmitted light and GFP fluorescence channels. The inlet port was connected to a 140 mL syringe (Covidien Monoject Syringe) and PTFE tubing (Cole Palmer PTFE#24 AWG tubing) with LB media with Kanamycin, 0.075% Tween-20, and 50 mM Methyl α-D-mannopyranoside. The waste port was connected to PTFE tubing and a 1L waste bottle. The height difference between the inlet and outlet was 39" corresponding to a flow rate of 11 mL/hr. Tween-20 and Methyl α-D-mannopyranoside were used in the media to prevent biofilming and therefore increase the longevity of microfluidic experiments. Tween-20 has been used by our group in many experiments without an adverse effect on *E. coli* [16, 18]. Methyl α-D-mannopyranoside inhibits normal surface attachment of type I pili but cannot be metabolized by *E. coli* [109]. After 16-24 hours of growth on LB media for cells to fill the traps, the media was switched to an HM9 minimal media described in Table 3.3, which was based on a previous study [110] and optimized for microfluidic *E. coli* growth with minimal traces of metals. Cells were grown on HM9 for 48 hours before inducing with heavy metals.

Heavy metal inductions occurred once a day for four hours. To induce, the HM9 in the inlet syringe was slowly pipetted out of the syringe and the HM9 + metal media was slowly pipetted in. To remove the heavy metal media after the 4 hour induction, the HM9 + metal media was pipetted out, and the remaining dead volume was washed with 2, 5 mL HM9 wash steps. Then the HM9 media was pipetted in to fill the syringe. The order of daily metal inductions was randomized for

each experiment. Inductions of each metal were performed across multiple experiments, with each experiment lasting 7-14 days. The number of inductions and concentrations for each metal is listed in Table 3.4.

Table 3.3: HM9 minimal media recipe.

Chemical	**Concentration (mM)**
Potassium Chloride	49.6
MOPS pH 7.2	40
β-Glycerol phosphate disodium salt pentahydrate	4
Dextrose	22.2
Ammonium chloride	18.70
Magnesium Sulfate, 7-Hydrate	0.2
Calcium Chloride Dihydrate	0.01
Iron(III) Chloride hexahydrate	0.001
Kanamycin sulfate from Streptomyces kanamyceticus	0.086
Tween 20	0.611
Methyl α-D-mannopyranoside	50

3.5.4 Live-cell imaging and data

Microfluidic devices were imaged using the Dynomics custom optical enclosure continuously every ten minutes in both the transmitted light and GFP fluorescence channels with a 1 second and 60 second exposure respectively.

Images were synced from the enclosure to a server via WiFi for further data processing. Custom software produced flat-field-corrected images in both channels in real-time to remove optical vignetting using the following equation:

$$C = m * \frac{R - D}{F' - D'} \tag{3.1}$$

where R is the raw image to be flat-field corrected (Figure 3.17a), D is the dark-current image for

Table 3.4: Metal induction count by concentration for fold change analysis and machine learning analysis

Metal	EPA MCL (ppb)	Concentration Tested (ppb)	Fold Change Induction Count	Machine Learning Induction Count
Arsenic	10	30	7	7
Antimony	6	1000	7	7
Cadmium	5	4.50	12	12
		8.99	0	1
Chromium	100	260	6	6
Copper	1300	127	8	8
Iron	300	279	6	6
		558	0	1
		5580	0	1
Lead	15	14.92	0	3
		29.84	6	6
		298.4	0	1
Mercury	2	2	6	6
Zinc	5000	50	5	5
		500	0	3
		1310	0	1

that device, taken at the same exposure settings as R, F' is a raw image taken by the camera with no device present, D' is the dark-current image taken at same exposure as F', m is the mean value for all values in the array $(F' - D')$, and C is the resulting corrected image (Figure 3.17b).

Flat-field corrected images were then registered to an extraction mask to account for the device drifting from thermal expansion, with mean transmitted light and green fluorescence channels from masked bulb and background regions extracted (Figure 3.18).

Raw data was initially processed as:

$$GFP_{Raw} = \frac{GFP_{Bulb} - GFP_{Background}}{GFP_{Background}} \tag{3.2}$$

For fold change analysis, fluorescence values were passed through a median filter (scipy.signal.medfilt, kernel_size=11) and normalized by promoterless strains (Figure 3.19). Promoterless strains U139 and U66 from the GFP promoter library [27] were spotted at various

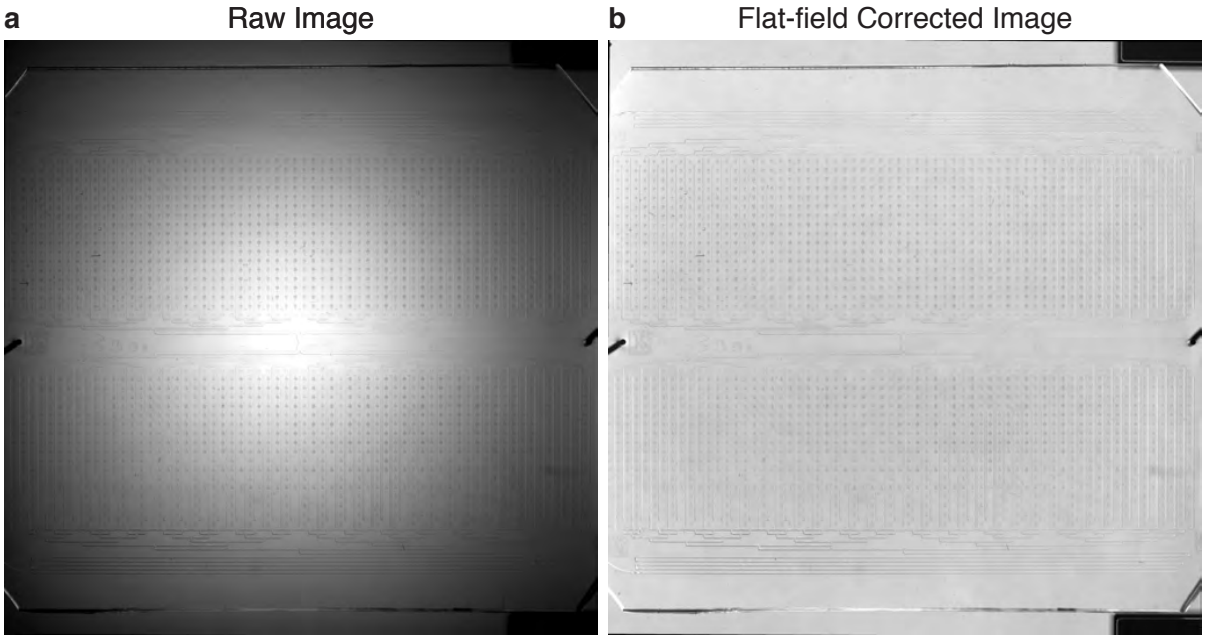

Figure 3.17: Flat-field correction with Dynomics. A full Dynomics image taken on custom optics before (a) and after (b) flat-field correction.

locations across each device, with 20 device positions of each, for 40 in total. Normalized fluorescence for Figure 3.6 was thus determined as:

$$GFP_{normalized} = \frac{scipy.signal.medfilt(GFP_{Raw}) - mean(scipy.signal.medfilt(GFP_{Promoterless}))}{mean(scipy.signal.medfilt(GFP_{Promoterless}))} \tag{3.3}$$

Fold change in Figure 3.6 was calculated as the quotient of the normalized fluorescence at the first and last time point of each metal exposure. P-values were determined by a dependent two-sided t-test on the log_2 fluorescent values at the start and end of each metal exposure (scipy.stats.ttest_rel).

3.5.5 Microfluidic device validation

In order to ensure that measured gene expression was independent of spatial position and to ensure responses weren't affected by previous inductions within a given experiment, a Dynomics chip was spotted with 2,176 replicates of the *zntA* promoter strain from the *E. coli*

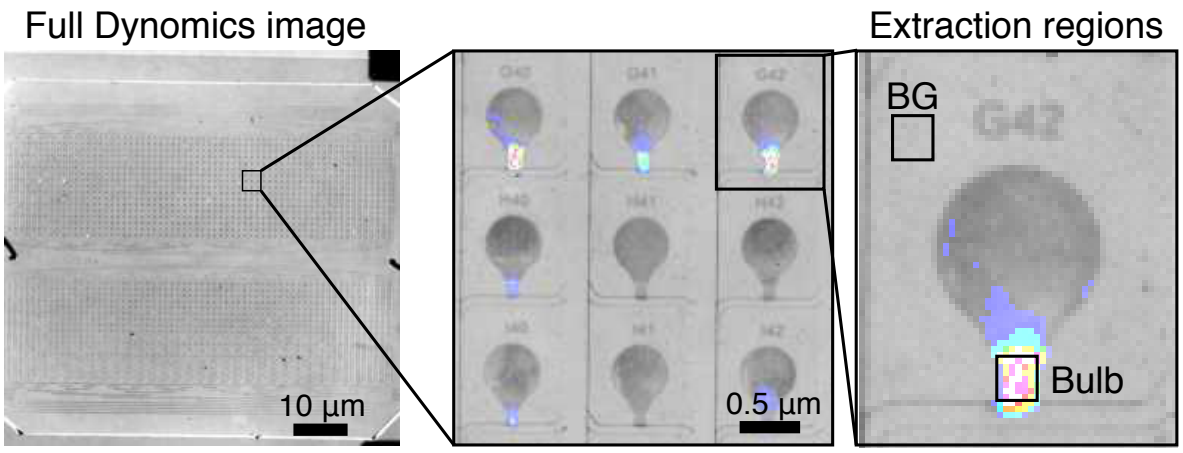

Figure 3.18: Data extraction from a Dynomics image. 2,176 device positions are simultaneously imaged in both transmitted light (gray) and green fluorescence (16 color) channels. At each time point for each device position, the mean fluorescence of the boxed bulb and background (BG) regions are extracted.

GFP-promoter library [27] and subjected to multiple replicate cadmium inductions. This strain is a reporter for the zinc and cadmium-responsive gene *zntA* and was chosen for its relevance in *E. coli* heavy metal metabolism and ability to be selectively activated via the heavy metal inducers used in the context of this work. As seen in Figure 3.20, strains across the device maintained statistically similar response parameters (amplitude of response and relaxation time) regardless of day of induction or spatial positioning on the device.

3.5.6 Gene ontology results for Figure 3.5

We performed enrichment analysis [111] on gene clusters obtained via agglomerative clustering of promoter activity, normalized between 0 and 1, of 1,995 strains responding to a 50 ppb Zn induction (Figure 3.5). Table 3.1 lists the enriched GO terms obtained for each cluster via `http://pantherdb.org/`. The statistical method used was Fisher's exact test with no correction for multiple testing. No correction was selected as we chose to minimize the number of false negatives over the number of false positives. We note that this increases the likelihood of finding enrichment terms by chance only. The reference list for the GO enrichment analysis was the

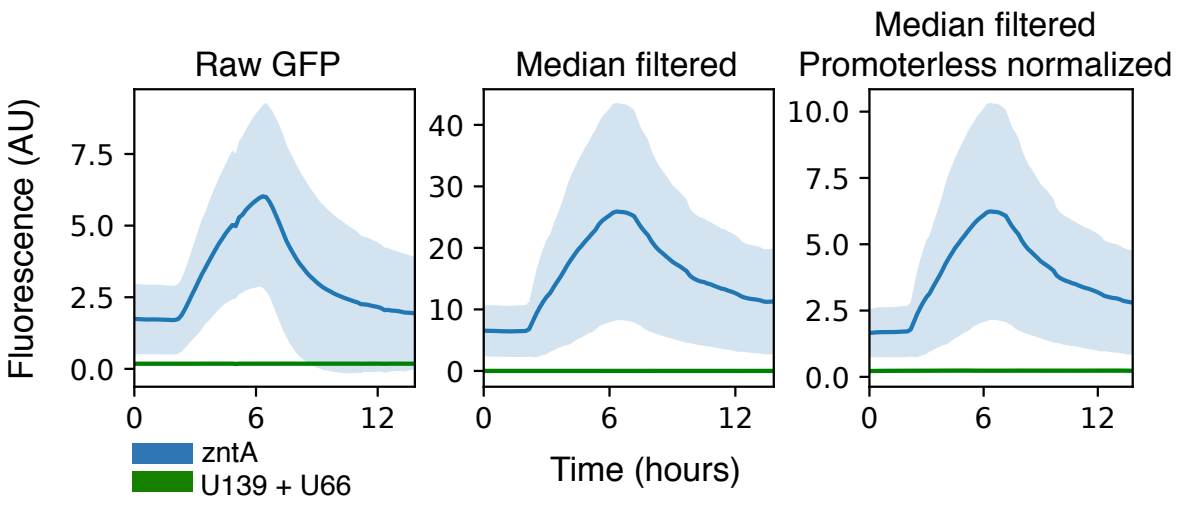

Figure 3.19: Normalization process for data presented in Figure 3.6. Raw data was median filtered, then divided by the mean traces of promoterless strains U139 and U66. Dark lines show the mean of all time series while shaded regions represent the standard deviation.

1,807 unique genes of the GFP-promoter library [27].

3.5.7 Machine learning

We transformed our eighteen standardized experiments' time points into a first derivative-based feature for the training and testing feature sets. All data used for machine learning results were pre-processed into first derivative-based features. Intuitively, a first derivative-based feature is an excellent candidate for any sort of machine learning model because it is what the human brain instinctively monitors when looking for changes in strain-promoter behavior. Any significant modification in the mean or variance of the first derivative of a given promoter while induced or uninduced could signify that the promoter is sensitive to that particular environmental perturbation. The engineering of a first derivative-approximation feature essentially amounts to distilling out the pure changes in the original feature's behavior, while effectively filtering out any changes that could be due to extraneous local or global environmental influences. Figure 3.21 illustrates the results of the feature engineering process.

In order to optimize the classifiers, extensive Bayesian Optimization searches were used

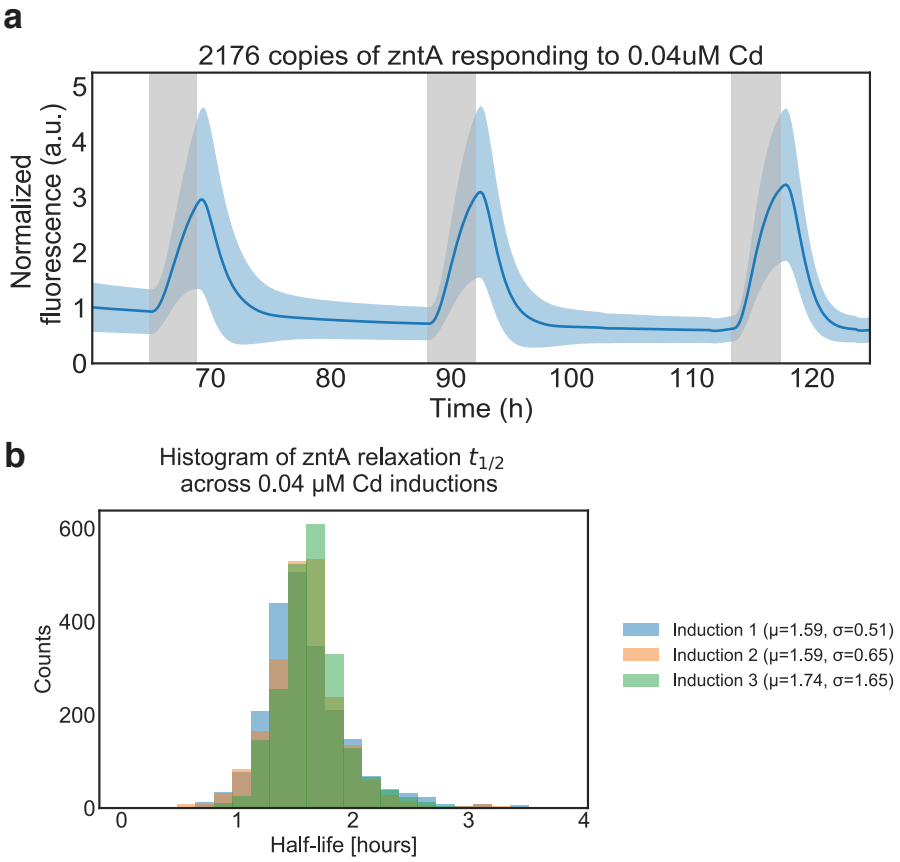

Figure 3.20: Device validation of a Dynomics chip. Device validation with a Dynomics chip with 2,176 replicates of the *zntA* strain from the *E. coli* GFP-promoter library [27]. a) Mean (dark blue) plus/minus one standard deviation (light blue) for the 2,176 strain responding to three consecutive four-hour 0.04 μM Cadmium inductions. b) Histograms of the decay half-life of post-induction response (i.e. the time taken for a strain's fluorescence levels to decay back down to their half-peak value) are overlaid for the three inductions plotted in a).

to find optimal hyperparameter combinations [112]. Throughout our hyperparameter searches, we used leave-one-out cross-validation on a per-experiment basis and appropriate overfitting-prevention strategies to ensure that any resultant classifier would generalize to future data sets. All classifiers were evaluated using the F_1-macro scoring metric. The F_1-macro score, which is the per-class average of the harmonic mean of precision and recall, was especially well-suited because of our data set's large multi-class imbalances, with water making up approximately 86% of the final feature set [113]. Finally, all generalization evaluations were performed by recording the results of using leave-one-out cross-validation with early stopping and then taking the mean

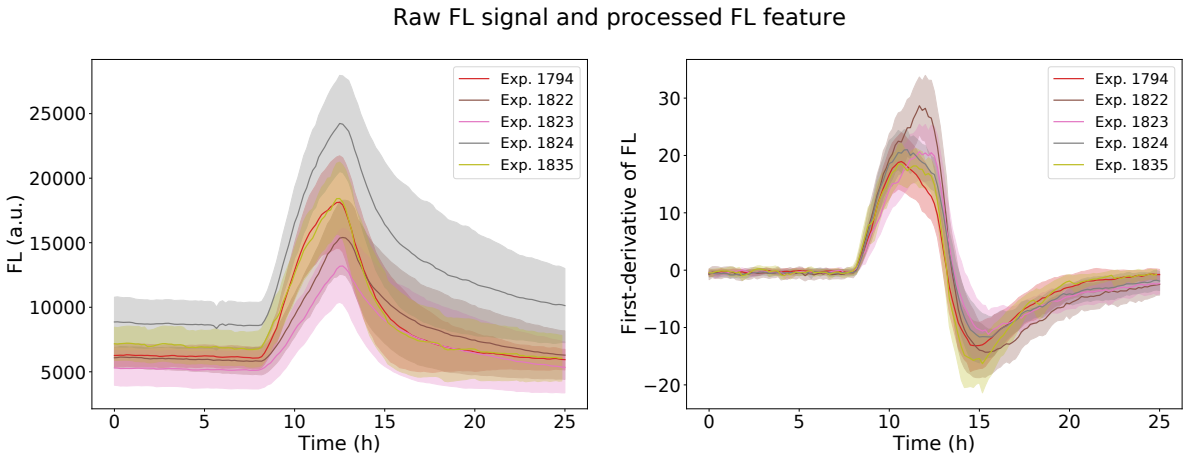

Figure 3.21: Raw and processed fluorescent signals. Our feature engineering eliminates a significant amount of intra-experiment variability by rendering the raw signal into a first derivative-like feature. This variability is due to differing hardware between Dynomics devices, among other sources.

prediction across the cross-validation's output.

Additional feature preprocessing

In addition to the data preprocessing and feature engineering that were explicitly enumerate in the previous subsection, other steps were often taken when training and testing any machine learning algorithms. These additional steps each dropped some portion of the final feature set, but *only after* the features had been calculated using the entire original experiments. Dropping these time points before calculating the features would have introduced potential discontinuities to any features approximating the first-derivative. All features were processed and cached in permanent memory. Only the cached features were used for any further analysis.

All experiments included transient periods over the first several days of the experiment. These transients were caused by the colonies' recoveries post-spotting shock, their growth-to-effluence within their individual traps, and their second recovering following the switch to minimal HM9 growth media. In addition, since most experiments were run until the microfluidic chips were deemed unusable due to clogging by biofilms, the final hours of most experiments did not

yield high-quality water data. Since the beginnings and the ends of the experiments represented non-steady state local and global growth conditions, the features from these periods were dropped prior to analysis and machine learning. This trimming was done by removing features more than eighteen hours before the first induction and more than four hours after the final metal induction.

In addition, inductions using undetectable metals, such as arsenic, mercury, and antimony, and special inductions that introduced any non-standard inducers were dropped from the feature set post-processing.

Mercury, arsenic, and antimony exposures were also fed into the classifiers, but despite extensive feature-engineering efforts, no classifier was able to successfully detect these metals. This negative result was most likely a consequence of the absence of the *E. coli ars* and *mer* operons from the Alon promoter library [114]. These operons are known to be highly efficient at sensing and exporting arsenic and mercury, respectively, from their cells [115, 116]. In addition, upon further review, we found that the tested concentrations of both metals were over an order of magnitude lower than the known in-batch minimum inhibitory concentrations for *E. coli* [117]. To focus classification on detectable metals, features during these metal inductions were relabeled as No Toxin.

3.5.8 Municipal water experimental set-up

Water samples were obtained from the Department of Water Management at the City of Chicago, in Chicago, IL, the Alex Orr Water Treatment Plant in Miami, FL, the New York City Department of Environmental Protection and Bureau of Water Supply in Corona, NY, the Seattle Public Utilities Water Quality Lab in Seattle WA, and the Alvarado Water Treatment Plant in San Diego, CA. HM9 media for each city water experiment was prepared by diluting 5x HM9 concentrate made from milliQ water with the water obtained from each city. The microfluidic device was initially grown on LB media with Kanamycin, 0.075% Tween-20, and 50 mM Methyl α-D-mannopyranoside until traps were filled to confluence and then switched to HM9 made with

city water for the remainder of the experiment. Cadmium diluted in the HM9 city water media was used to perform inductions as described in Section 3.5.3.

3.5.9 Gold King Mine spill experimental set-up

Water was collected from Mexican Hat, Utah in August, 2015 when the Gold King Mine Spill plume reached the collection point in the San Animas River. Samples were stored in 0.5% HCl Acid until tested. HM9 media was prepared by diluting 5x HM9 concentrate made from milliQ water with filtered San Animas samples. The pH was adjusted to 7.05. The metal concentrations of the HM9 San Animas samples were tested by ICP-MS at the Environmental and Complex Analysis Laboratory (ECAL) at UC San Diego. Four hour inductions were performed as described in Section 3.5.3.

3.6 Acknowledgements

This chapter contains material under review at the *Proceedings of the National Academy of Sciences* as Genome-scale transcriptional dynamics and environmental biosensing. Graham, Garrett*, Csicsery, Nicholas*, Stasiowski, Elizabeth*, Thouvenin, Gregoire*, Mather, William H., Ferry, Michael, Cookson, Scott, and Hasty, Jeff. (*equal contribution) The dissertation author was one of the primary authors and researchers of this material.

Chapter 4

DynaScreen: A multiplexed-microfluidic platform for the dynamic phenotypic screening of synthetic gene circuits

4.1 Introduction

4.1.1 The need for continuous phenotypic screening systems

As described in Chapter 1, synthetic biology strives to achieve the forward engineering of biological systems by piecing together DNA modules to build genetic circuits with complex functions [4, 5]. However, in practice, forward engineering of gene circuits requires frequent design-build-test-learn cycles, often with little quantitative guidance as to which genetic modules will best produce the desired output. Fortunately for synthetic biologists, a vast number of genetic modules and strains exist within the scientific community and in resources such as the Registry of Standard Biological Parts [4]. These libraries and registries of genetic parts enable the rapid prototyping of many circuit variants, however, few of these modules are well-characterized, and for those that are, it is primarily done with batch characterization in micro-titer plates. This poses

a challenge to forward-engineering as future gene circuits are deployed in real-world settings with temporally fluctuating environments, such as bioreactors [118, 119], living systems [9, 120, 121], or nature [122]. Furthermore, over a decade of research from our group and others has revealed that batch characterization is not sufficient to predict and model gene circuit function [21, 123] or to predict behavior in a continuous system [6, 9, 17, 18]. In this chapter, I describe the adaptation and application of the previously described Dynomics platform (Chapters 2 and 3) to dynamically screen libraries of synthetic gene circuits, biosensors and oscillators, under changing environmental conditions.

4.1.2 Existing approaches for phenotypic screening

Phenotypic screening is widely used in drug discovery, genomics, and biological research including synthetic biology. In all of these cases, screening is typically performed in high-throughput, batch systems. One of the most common methods of phenotypic screening is those using microtiter plates [124, 125]. These methods utilize low volumes (200 μL or less), offer the capability to test anywhere from 96-1,536 unique conditions, and can be tied to many readout types including fluorescence, luminescence, and optical density. Similarly, agar plates have been used to screen solid colony phenotypes with up to 6,144 colonies one a single plate [126]. With agar-plate-based methods, assay readout is similar to liquid-plate assays, the number of strains that can be screened is increased four-fold or greater, but only one environmental condition can be tested per plate. In other cases, flow-cytometry has been used as an alternative to plate based-screening [127]. The aforementioned approaches are widely used throughout academia and industry and have brought forth great advances in biological understanding, drug development, and genetic engineering. However, the methods described here are based on batch growth and allow little-to-no temporal control of the growth environment. Towards the high-throughput screening under continuous growth has been the development of the various multiplexed microfluidic platforms described in Chapter 2 [24, 30, 50, 128]. However, the potential of these devices to

continuously screen libraries of synthetic circuits has not been fully explored.

4.1.3 Methods for generating gene circuit libraries

Within synthetic biology, generating a variety of gene circuits can be accomplished in several of ways. As described and utilized in Chapters 2 and 3, libraries of microbial strains can be created by changing a specific element in each library member such as the promoter driving GFP [27] or the protein tagged with a fluorescent protein [26, 52]. Resources such as the Registry of Standard Biological Parts can be especially useful here, as typical gene circuit troubleshooting will involve prototyping with several previously used genetic modules. In less defined cases, or when parts delivering the desired function aren't suspected to exist, mutant libraries can be generated through either error-prone PCR or site-directed mutagenesis of genetic elements and protein coding sequences [129, 130]. Lastly, in cases where desired phenotype can be tied to growth rate, adaptive laboratory evolution and competition experiments can be used evolve and screen strains based on growth-characteristics [131].

4.2 An acoustic loaded microfluidic platform for dynamical gene circuit screening

4.2.1 Limitations of contact spotting from agar plates

To develop a multiplexed device for dynamical screening of numerous bacterial strains in parallel, we identified minor shortcomings in the platform developed in Chapter 2 and deployed in Chapter 3. While incredibly robust when screening the same arrangement of two-thousand strains against multiple environmental conditions, as was the case in Chapter 3, re-arraying the strain layout on a device is slow, taking up to a full week, and not conducive to screening hundreds to thousands of strains with a high layout turnover (Figure 4.1).

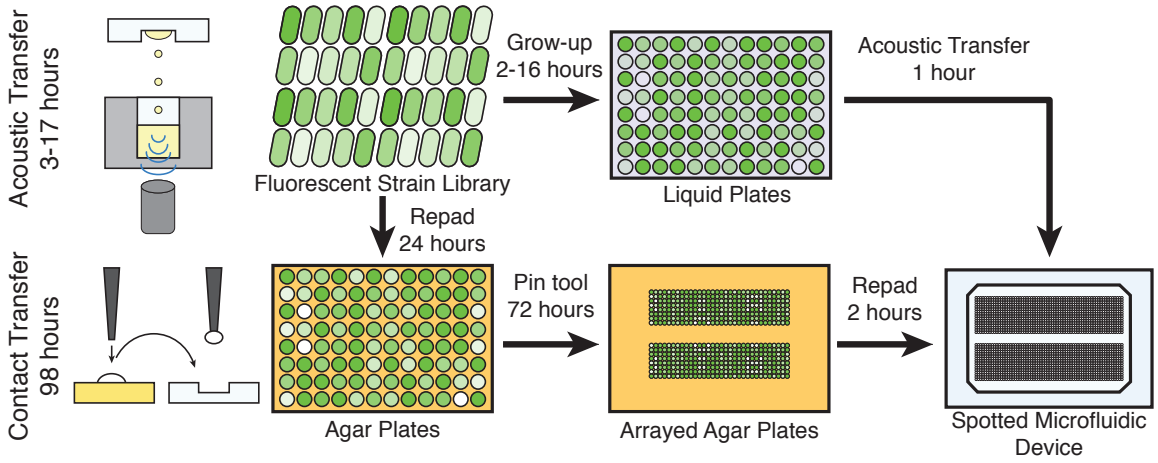

Figure 4.1: Experimental throughput of contact and acoustic cell transfer. For contact transfer from agar plates to PDMS devices, cell re-arraying takes roughly 96 hours for a 2,000 strain device and an additional 2 hours for transfer from agar to PDMS. Acoustic transfer does not require the pre-rearraying as the Echo liquid handler can eject roughly 60 drops per second, allowing for on-the-fly arraying. Cells are grown to a desired OD on liquid plates, typically 2-16 hours, and transferred directly to a spotted microfluidic device.

To improve throughput, we looked towards acoustic droplet ejection [132] which could enable us to rapidly and directly transfer overnight bacterial liquid cultures to PDMS devices. While this would not significantly reduce day-of-experimental spotting time (2 to 1 hours), it would eliminate the need for re-arraying large libraries, potentially saving a full work week of re-arraying for each unique device layout. This is made possible as the Singer Stinger transfers 190 colonies a minute while a commercial acoustic liquid transfer robot, the Labcyte Echo, can transfer 3,600 droplets a minute.

Beyond limited experimental throughput, contact spotting also limits the types of experiments that can be performed with this loading method. While excellently capturing dynamics of a first-order-like system such as the inducible promoters in the *E. coli* promoter library [27], second-order systems such as genetic oscillators perform poorly. In the case of a previously developed lysis oscillator [9], cells spontaneously lyse and express GFP in synchrony once the presence of a quorum-sensing molecule AHL reaches a certain threshold. However, due to differences in growth rate and small molecule diffusion between the front and back of the large

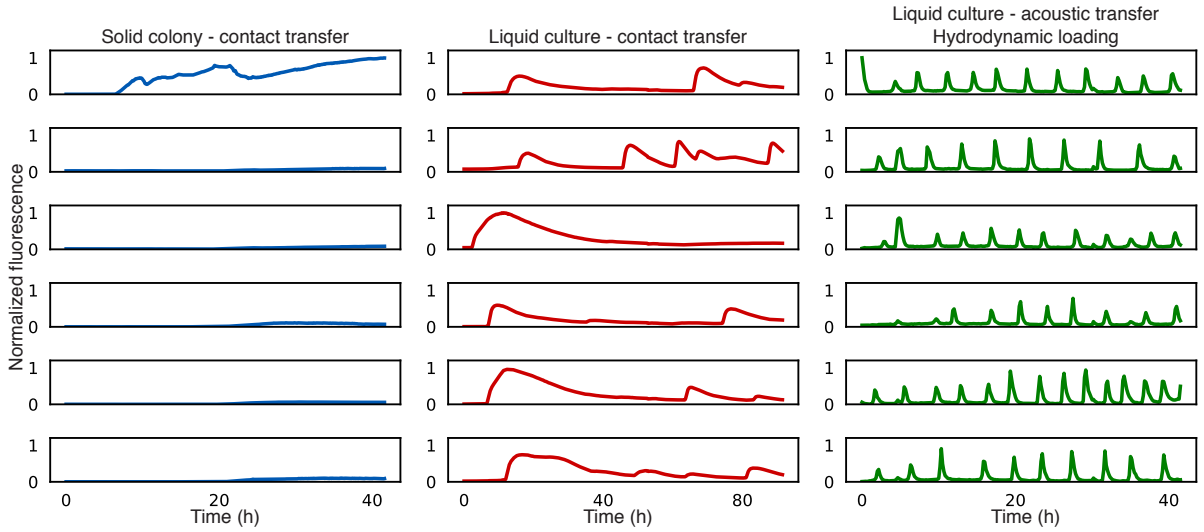

Figure 4.2: Synchronized lysis oscillator dynamics with different spotting methods. Fluorescence traces from six cell trapping regions are shown for each method. ROTOR spotting from solid colonies leads to some oscillatory behavior, though widely inconsistent between traps. Hand-spotting liquid cultures to the same trap geometry enables visible oscillations in all cell traps, however, oscillation dynamics are inconsistent both temporally and between cell traps. Loading multiplexed devices from liquid culture and allowing hydrodynamic loading of downstream cell traps enables robust and consistent oscillations across all cell traps.

spotting area cells are unable to consistently lyse in synchrony (Figure 4.2). As is apparent in fluorescent traces from each spotting method, synchronized lysis is further hindered by the immovable mass of cells and agar from contact spotting of solid colonies, often resulting in no synchronized lysis at all.

We sought to instead load liquid cultures onto the microfluidic devices since synthetic lysis oscillators are typically loaded fluidically onto microfluidics during early exponential growth before cell lysis has occurred [9], a growth condition that cannot be replicated on agar plates. Hand-spotting of liquid cultures produces better oscillator behavior but presents two challenges. First, oscillations are still inconsistent due to the large geometry of the cell-spotting and trapping region. And second, hand-spotting hundreds to thousands of strains is not logistically feasible. Thus, we sought to construct a new automated protocol and device design for liquid loading of cells across the Dynomics platform. Here, I describe the development of a pipeline for the direct deposition of *E. coli* onto a microfluidic device with an acoustic liquid handler, while maintaining

cell viability and gene circuit function.

4.2.2 Microfluidic redesign for acoustic cell loading

A microfluidic array was developed for the continuous culturing and monitoring of *E. coli* strains loaded by acoustic droplet ejection using a Labcyte Echo liquid handler. This system, capable of co-culturing 144 distinct strains, consists of a single inlet branching over a manifold of cell spotting and trapping regions before recombining into a single outlet (Figure 4.3A). Each spotting region has built-in error tolerance for imprecision in acoustic transfer equipment (explored in depth in Figure 4.4), allowing cells to be acoustically spotted within, or overlapping with, the spotting features while maintaining reliable cell trap filling. Upon device setup, growth media from the manifold channels flows through the spotting region features carrying spotted cells to the downstream cell traps (Figure 4.3B-C). Complete washout of the spotting region is prevented by the presence of low-flow internal channel, observed experimentally and in finite element analysis modeling in COMSOL (Figure 4.5). Downstream cell traps are hydrodynamically loaded when empty, as cells carried by flow through the traps are stuck in the 0.5 μm layer. Within roughly 8 hours, cell traps reach confluence, preventing further flowthrough.

With this system, cells can be loaded directly from liquid culture, grown under a precise and dynamic environment, and measured continuously with experimental run times lasting 1-2 weeks. Both biosensor and genetic oscillator circuits were tested in this device with acoustic loading, imaged in the down-stream hydrodynamically loaded cell traps, and performed as they would in traditional fluidically loaded microfluidic experiments (Figure 4.6). Biosensor circuits responded with increased fluorescence with the addition of inducer molecules and with decreased fluorescence upon their removal. Fluctuations of transmitted light and fluorescence were measured over several days in synthetic lysis oscillators [9]. Improved oscillator performance in this device is attributed to a healthier cell state when grown in liquid culture and a vast decrease in the cell trap size, from a previously described 500 μm diameter spotting region and cell trap to a 70 μm

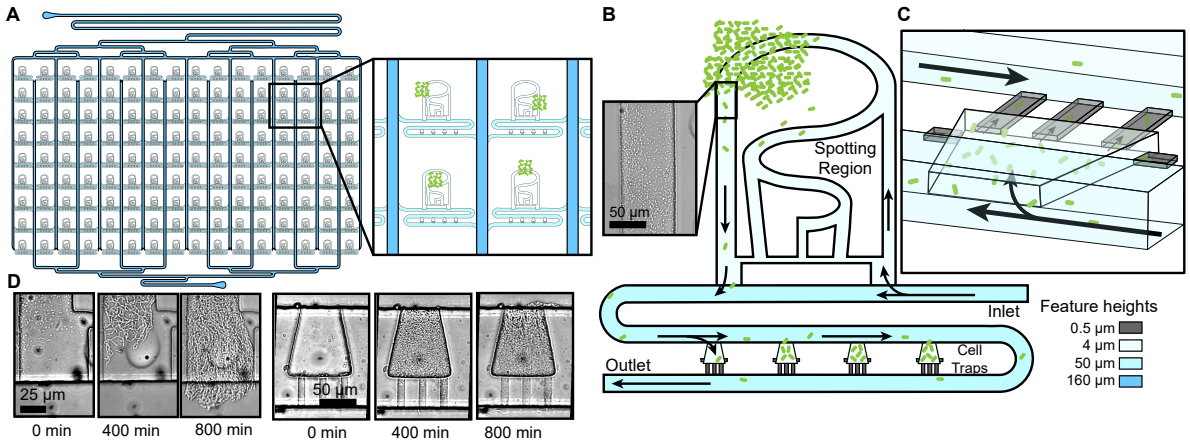

Figure 4.3: An acoustic droplet ejection compatible microfluidic device. (A) A microfluidic device was developed with 144 individually spottable cell growth regions. The inset shows typical spatial variability of 2.5-5 nL *E. coli* droplets. (B) Each cell growth region is comprised of a spotting region at which cells are deposited via acoustic liquid transfer. (C) Cells flow out of the spotting reservoir and into the cell trapping region where they are imaged. When traps are empty, media flow will carry cells into the cell traps, where they get caught in the 0.5 µm layer and grow to confluence. Once cell traps fill, flow through the traps is blocked and media flow proceeds perpendicular to the traps in the main channel. (D) *E. coli* growth in the spotting region is slow with sick, elongated cells, while growth in the cell traps is faster with healthier cells.

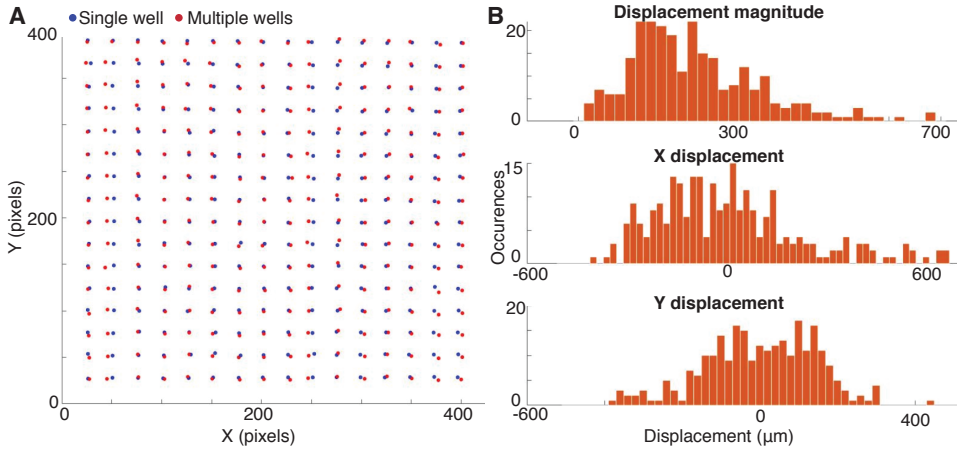

Figure 4.4: Acoustic transfer imprecision. (A) 256 cell dyed cell culture were transferred to filter paper, both from a single source well and from multiple source wells. Spot positions were determined in ImageJ. (B) The difference between each cell spot was determined and the total displacement, as well as displacement in x and y was determined. The average deviation from the spot grid was 224 µm, while the most extreme observed deviation was 682 µm.

wide downstream cell trap (Figure 4.2).

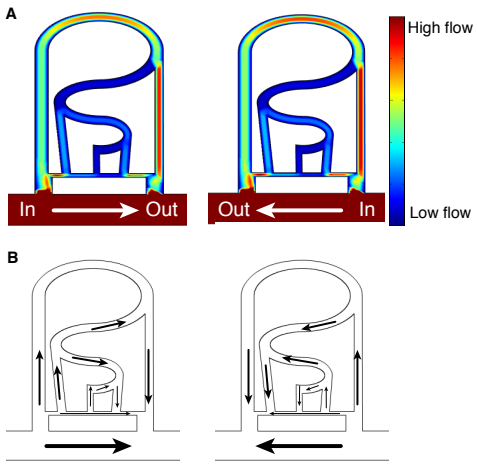

Figure 4.5: COMSOL modeling of cell spotting region. (A) Flow rate throughout the cell spotting region. High flow is observed in the main channel, while low flow is observed in winding side channels. (B) Flow direction throughout the cell spotting region. Arrow size not to scale with flow rate.

4.3 High-throughput dynamic screening of synthetic gene circuits

4.3.1 Development of fluorescent biosensor strains

As described in Chapter 3, whole-cell biosensing applications have been a forefront pursuit of synthetic biologists since the advent of the field [13, 59]. A prominent area of biosensor development has been environmental biosensors to determine the safety and quality of water, both for drinking and agricultural applications. Though environmental sensors have begun to approach real-world deployability, many such sensors exist in batch systems [128, 133–135] which are often limited to a single measurement of a single environmental sample or compound. Furthermore, reporting modules are most commonly fluorescent or luminescent proteins [18, 29, 59, 60, 128, 133–135], which produce a strong signal but can require specialized tools to measure.

Here, I present multiple tools to expand the utility and deployability of synthetic biosensors. First, I illustrate how the multiplexed, acoustic loadable microfluidic platform enables

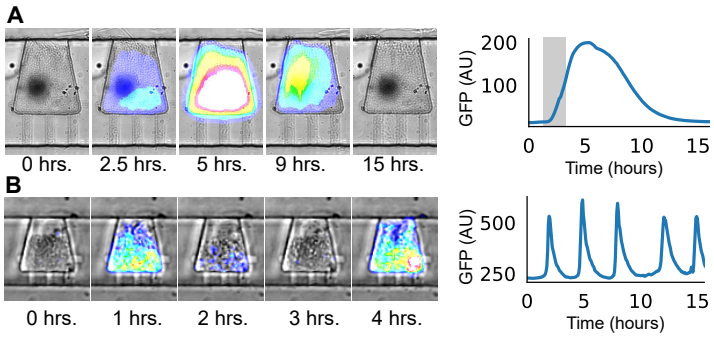

Figure 4.6: Demonstration of cellular dynamics which can be measured in this device. (A) A simple inducible circuit driving GFP is exposed to a two hour exposure to an inducer compound (1 μm Cd, gray bar). (B) A free-running synthetic oscillator driving the expression of GFP and lysis protein at regular intervals.

continuous measurement of the response of synthetic biosensor strains to dynamic environmental perturbations on week timescales. We demonstrate that the co-culture of these sensor strains enables the specific real-time detection of five heavy metals of interest: arsenic, cadmium, lead, mercury, and copper. Second, we demonstrate that with the substitution of a fluorescent reporter with a protein driving programmed cell lysis, the reporting mechanism of existing fluorescent biosensor strains can easily be converted to one based on optical density or electrical impedance. We show that such circuit redesign eliminates the reliance on fluorescence-based optics, and in some cases, increases strain sensitivity. Finally, we show the dual-utility of our acoustic loadable microfluidic platform not only for biosensing applications, but for the screening of synthetic strain libraries. Such screens are used to expand the set of sensors with both fluorescent and lysis-based reporting for better metal detection.

This platform was initially deployed for the monitoring of five synthetic *E. coli* biosensor strains designed to sense copper, arsenic, cadmium, lead, and mercury (Fig. 4.7A). Each sensor construct consists of a metal inducible promoter driving the expression GFP, and in some cases, its own regulating transcription factor. The *E. coli* MG1655 chassis strain genome was left unaltered and in all but the mercury sensing strain contains natively expressed elements which are known to interact with the engineered plasmids.

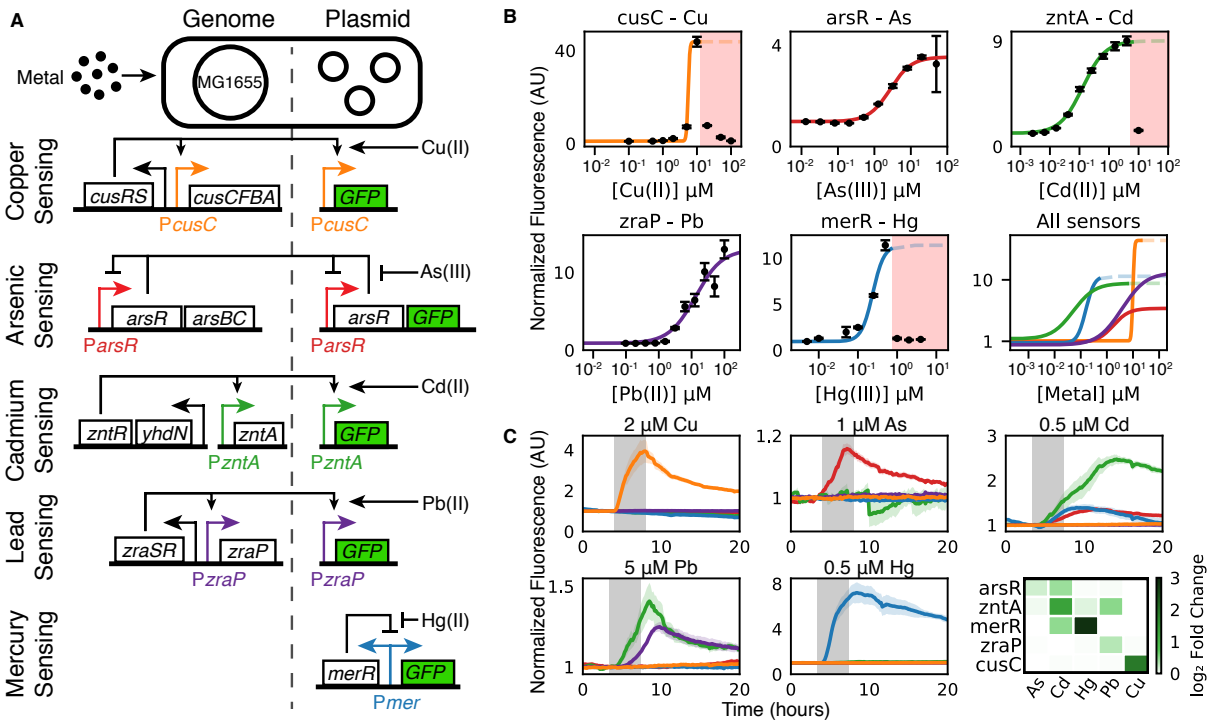

Figure 4.7: Synthetic fluorescent biosensor strains for the detection of heavy metals. (A) Five plasmids were engineered to respond to five metals (Cu, As, Cd, Pb, Hg) in the MG1655 strain of *E. coli*, with circuit design presented at the right of the vertical dashed line. Gene circuit architecture to the left of the vertical dashed line represent native elements of the MG1655 genome that interact with the engineered plasmids. (B) The metal-fluorescence dose-response was measured for each of the biosensor strains. Concentrations in red affect cell growth and were not used to fit the dose-response curve. (C) The fluorescence response of the five sensor strains as measured in the acoustic loaded microfluidic device. All strains were subjected to 4 hour exposures of each metal. Solid line represents the mean of 4 cell traps and shaded region represents standard deviation.

The native *E. coli* ArsR promoter was selected for the arsenic sensing strain. 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 [136]. 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 [137]. 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: *zntA* [138] for cadmium and lead, *cusC* [139] for copper, and *zraP* [140] for lead. All the biosensor constructs described previously were placed on a medium-copy p15A origin plasmid with a kanamycin resistance gene. All sensor strains were tested in batch by measuring the dose-response curve (Fig. 4.7B). Because of the toxic nature of heavy metals, several sensors arrest growth before the GFP-toxin response is saturated, however, a detectable level of fluorescence is measured at non-toxic concentrations for all strains. The performance of these sensors strains was evaluated in acoustic loaded devices, enabling sequential exposures of all strains to multiple metals. 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 media can be uniquely identified (Fig. 4.7C).

4.3.2 Development of lysis-based sensor strains

As an alternative to fluorescence-based methods of reporting in biosensor strains, we tested the viability of optical density-based reporting with a simple swap of the fluorescent reporter gene (GFP) with the E lysis protein from bacteriophage ϕX174 using Gibson assembly (Figure 4.8A) [141]. For initial testing, each strain was grown in a 1 mL HM9 batch culture with and without its respective inducing toxin. All strains except *zntA* exhibited inhibited cell growth, visible by eye, when grown under the constant presence of its respective inducer compound (Figure 4.8B).

The strains were then further investigated in acoustic loaded microfluidics. I observed that that these strains were effective in continuous culture, with induced lysis in the presence of inducing toxin, and regrowth with the removal of the toxin, (Figure 4.8C-D). When all strains were grown together and subjected to induction series of each metal of interest, little cross-talk was observed between the metals (Figure 4.8E and Figure 4.9). Signal from each cell trap was measured as the normalized lysis rate, or the first derivative of TL, normalized between zero and

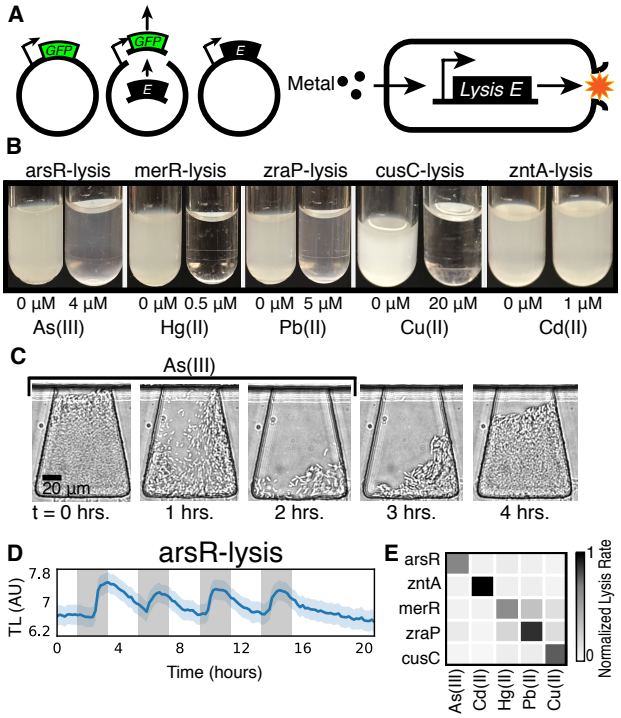

Figure 4.8: Lysis-based reporting for synthetic metal sensor strains. (A) The GFP reporter gene was replaced with the E lysis protein from phage X174 in the five sensor strains to create a lysis-based method of reporting metal presence. (B) 1 mL of cell culture in HM9 grown with and without the respective inducing toxin. (C) In microfluidics lysis sensors lyse and empty cells traps in the presence of induce signal, and grow and refill in its absence. (D) *arsR* lysis-based sensor responding to 2 hour pulses of 1 µm As(III) (gray bars). Solid line represents the mean of multiple cell traps and shaded region represents standard deviation. (E) First derivative of TL of each lysis strain measured in microfluidic experiments. Metal concentrations for all measurements were: 1 µm As(III), 0.5 µm Cd(II), 0.5 µm Hg(II), 5 µm Pb(II), and 2 µm Cu(II)

one for all strains. In general, while the lysis-based sensors are incredibly responsive upon the exposure to a metal, cell recovery back to full, growing cell traps can take long periods of time. In the case of *cusC*, cell traps take nearly four days to refill. With fluorescence strains, all sensors return to baseline levels of fluorescence within 24 hours of a 4 hour metal exposure. Therefore, while the lysis-based sensor strains may be used for the real-time detection of heavy metals, best performance would consist of the co-deployment of these strains with fluorescence-based sensors.

When exposing cell culture in early exponential phase to serial dilutions of each metal, we observed that all five fluorescent biosensor strains exhibited induced lysis in the presence

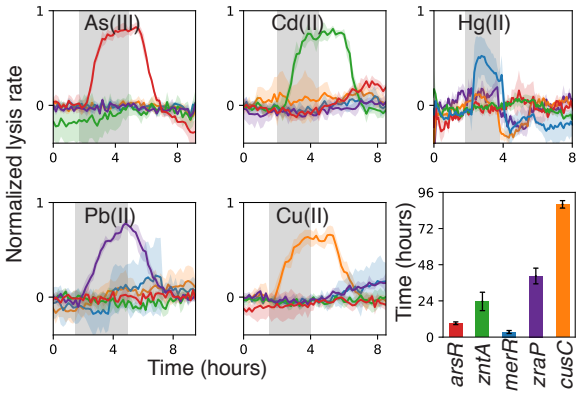

Figure 4.9: Lysis-based reporting in microfluidics. Normalized lysis rate of the lysis-based sensors when exposed to 1-2 hour pulses of each metal (gray bars). Solid line represents the mean of four cell traps, and shaded region represents standard deviation.

of their respective metal, however, many (notably *cusC* and *zntA*) were able to recover, at least partially, and grow after the initial lysis event. Dose-response curves were measured for each of the lysis-based reporters, quantifying the magnitude of the lysis event immediately after the introduction of toxin and 8 hours after the introduction of toxin as well as the rate at which cells lysed. In both cases the metal concentration necessary to induce half of the maximum lysis (EC_{50}) was lower than that required for fluorescence-based reporters (Table 4.1). This is likely due to the destructive nature of the lysis protein, resulting in switch-like, "all-or-nothing" lysis events.

Table 4.1: Dose-response EC_{50} of lysis-based sensors (ppb)

Strain	**GFP**	**Lysis mag. (1 h)**	**Lysis mag. (8 h)**	**Lysis rate**
cusC	347	40	-	35
arsR	196	87	95	91
zntA	15	7.5	-	5
zraP	2850	178	143	124
merR	48	16	24	20

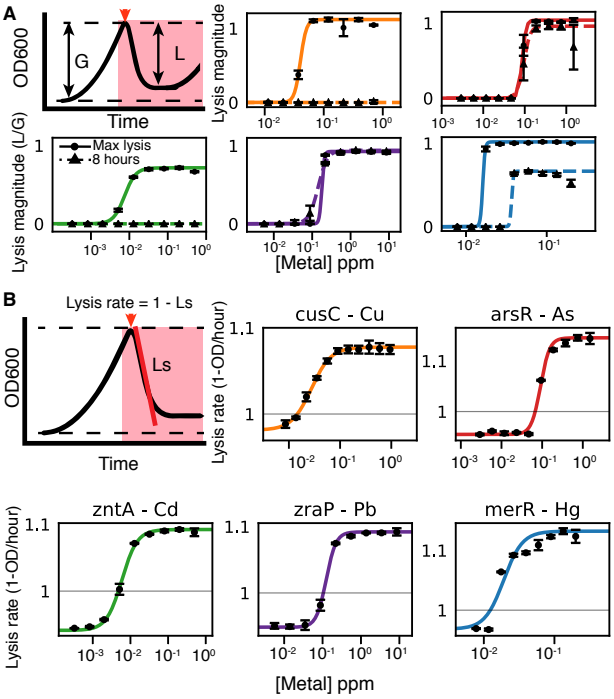

Figure 4.10: Induction curves of the lysis-based sensor strains. (A) The magnitude of lysis (left) is measured as the ratio of the change in OD due to lysis over the change in OD due to growth before toxin exposure (red arrow). Lysis magnitude is measured at the initial lysis event (solid line, black circles) and 8 hours after toxin exposure (dashed line, black triangles). Error bars represent standard deviation. (B) The lysis rate is represented as the difference of 1 and the minimum first derivative after toxin exposure. Gray line at lysis rate of 1 indicates an OD derivative of 0. Error bars represent standard deviation.

4.3.3 Electrochemical measurement of lysis-based sensors

A substantial challenge in synthetic biology is minimally tracking gene expression without the needed for fluorescent proteins and the associated imaging equipment. As a proof of concept of the utility of lysis-based reporting mechanisms, we measured the impedance of cell culture in a chemostat during a lysis event. We observed that when induced lysis occurred, both turbidity and admittance of the culture decreased (Figure 4.11). This paves the way for the biosensing of heavy metals without the need for any optics, only electrical probes. The elimination of bulky optics can significantly reduce the size and cost of future sensing platforms.

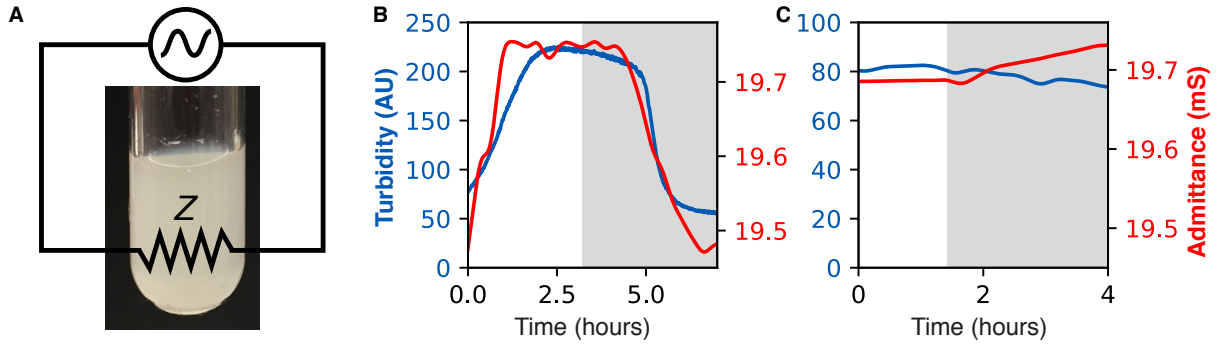

Figure 4.11: Interfacing bacterial circuits for electrochemical sensing. (A) A conceptual figure representing a custom electrochemostat used to measure the impedance of a bacterial culture. (B) Profiles of the admittance (red) and turbidity (blue) using an arsenic-sensitive strain in an electrochemostat device. The shaded region represents the duration of 250 ppb arsenic induction. (C) Profiles for the As-lysis strain induced with 250 ppb copper.

Table 4.2: ICP-MS measurements of metals water sources (ppb). * indicates national secondary standard, a non-enforceable federal guideline. Bolded values exceed Environmental protection agency (EPA) federal standard for drinking water.

	Water treatment	**Suburban**	**Ocean**	**Apple Juice**	**HM9**	**EPA (MCL)**
Cu	64.28	590.97	42.86	**1542.10**	8.29	1300
As	6.34	4.20	13.28	**87.12**	ND	10
Cd	0.25	0.21	0.52	1.27	0.32	5
Pb	5.91	1.14	ND	7.58	**30.68**	15
Hg	ND	ND	ND	ND	ND	2
Zn	92.86	48.03	ND	1469.68	20.53	5000*
Fe	273.84	41.60	**408.31**	**12469.97**	**483.11**	300*
Cr	23.69	1.17	20.81	89.44	2.55	100

4.3.4 Screening fluorescent biosensor strains under multiple media conditions.

To demonstrate the screening capabilities of the acoustic-loadable platform, and to evaluate the potential of these strains for the real-time detection of heavy metals on real environmental samples, the sensor strains were subjected to 6 consecutive exposures of metals on four different water supplies: (1) water from the Alvarado Water Treatment Plant in San Diego, CA, (2) suburban tap water from San Diego, CA, (3) Ocean water from Black's Beach in La Jolla, CA, and (4) apple juice (Fig. 4.12). The four conditions were tested in four separate chips in parallel on a

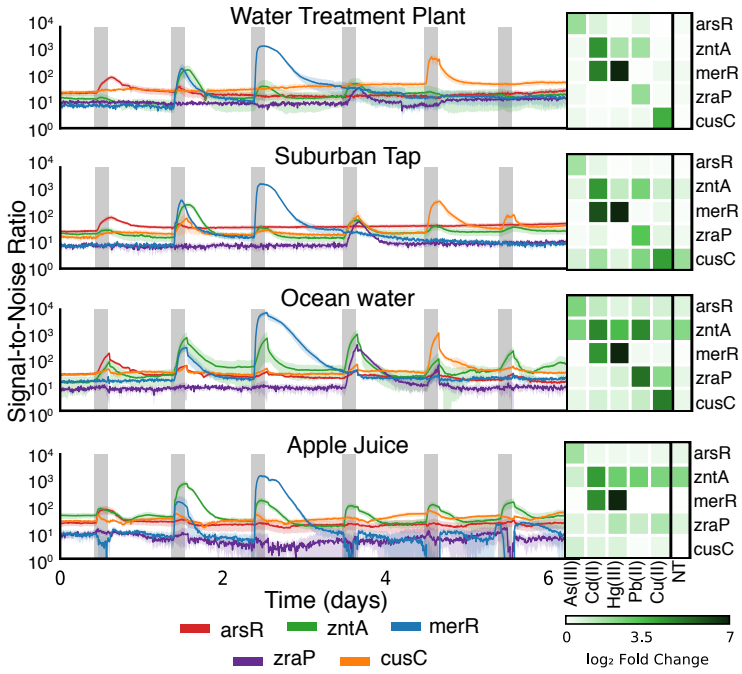

Figure 4.12: Biosensor performance on different water sources. The five fluorescent sensor strains were tested on four environmental water sources: water from a treatment plant, suburban tap water, ocean water, and commercial apple juice. Metal exposures were performed daily (gray bars) at the following concentrations: 1 μm As(III), 0.5 μm Cd(II), 0.5 μm Hg(II), 5 μm Pb(II), and 2 μm Cu(II). No toxin (NT) represents an exposure to the water sample with no additional dissolved metals. Mean signal-to-noise ratio is represented as a solid line and standard deviation as the shaded region.

single imaging platform. In additional to screening on the biosensor strains, these environmental samples were also subjected to ICP-MS to determine specific quantities of trace metals in each media (Table 4.2).

All sensor strains performed nearly identically to lab samples when tested in water from a water treatment facility. On suburban tap water, sensors can readily distinguish between different metals when tested on suburban tap water. However, the presence of copper in this water source contributes to a baseline response in the *cueO* strain upon each metal exposure. Ocean water induces an unconventional response in the sensor strains, likely due to cell stress in the high-salinity environment. These responses exhibit discontinuous relaxation dynamics upon the removal of ocean water, and more off-target excitation in the presence of each metal. Most notably,

zntA is activated in the absence of any added metal, despite low levels of trace cadmium and zinc in ocean water. Finally, apple juice proved to be a difficult sensing background, especially with its high autofluorescence masking the sensor response. The *zntA* sensor is continuously activated by apple juice, likely due to the high amount of trace metals, particularly zinc which is known to interact with the ZntR regulator [138]. Furthermore, copper and lead are not detected on this background media. This could be due to the already near-toxic to *E. coli* levels of trace copper in apple juice, possible chelating agents in apple juice, or the high background fluorescence masking the signal. Nonetheless, these results demonstrate the utility of this microfluidic platform to rapidly determine on which background medias these sensors can reliably report the presence of different metals.

4.3.5 Screening of libraries of synthetic gene circuit mutants.

Libraries of the *cusC* and *zntA* fluorescent sensor strains were created through error-prone PCR and subsequent cloning of the promoter region on each sensor plasmid. Strains were loaded onto 144- or 48-strain microfluidic devices and subjected to multiple exposures of either Cu(II) or Cd(II) at different concentrations (Fig. 4.13A,E). The first derivative of fluorescence and the time constant, τ, were measured for the time series data as representative dynamic parameters for differentiating the on- and off-dynamics of the responsive strains (Fig. 4.13B,C,F). Among the *cusC* variants, strains with both increased and decreased sensitivity were identified, denoted as *cusC*_9 and *cusC*_12 respectively (Fig. 4.13B). With significantly varying response dynamics in response to copper introduction, the co-culture of these variant strains on a single device expands the sensors sensitivity and dynamic range, while further serving as candidates for "traffic light" sensors, for which less external calibration is needed to discern copper concentration [142]. Furthermore, when the increased sensitivity *cusC*_9 and decreased sensitivity *zntA*_4 strains were cloned into their respective lysis sensors, we see an improved lysis stability as compared to those presented in Figure 4.8 (Fig. 4.13D,H).

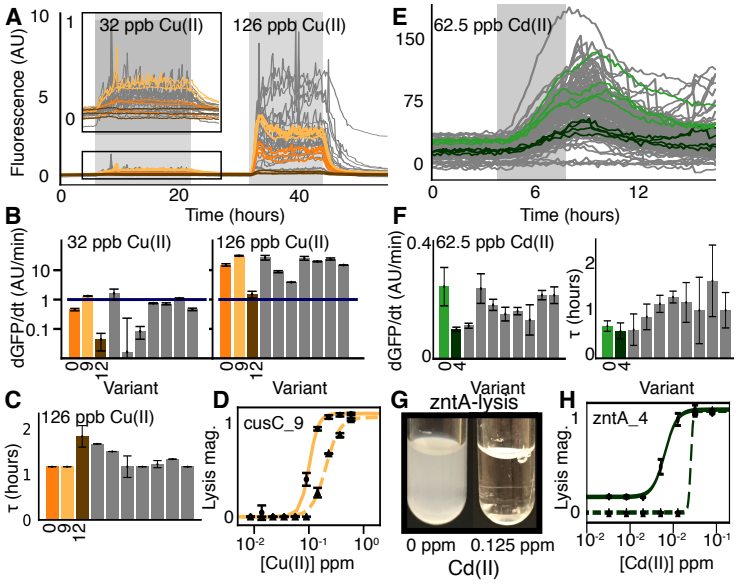

Figure 4.13: Dynamic library screening in microfluidics. (A) 16 mutant *cusC* strains in triplicate subjected to low and high concentrations of Cu(II) in microfluidics. Wild-type strain (orange), a strongly-expressed variant (yellow) and a weakly-expressed variant (brown) are highlighted. (B) Extraction of the first derivative for each strain (10 shown) reveals a range of responsiveness at each metal concentration. With a threshold detection threshold defined at a first derivative of 100/min, the yellow variant is the only above-threshold strain at 32 ppb Cu(II), while all three highlighted strains are above the threshold at 126 ppb. (C) Dynamic screening allows us to better understand how these strains will perform *in situ*, with an important parameter being the time constant τ representing the time to reach the e fraction of the maximum amplitude. (D) Lysis performance of the *cusC_9* variant lysis strain in batch, with lysis magnitude at 1 hour (solid line, circles) and 8 hours (dashed line, triangles) shown. (E) Response dynamics were measured for 16 *zntA* variant strains under a single toxin concentration. (F) Similar response dynamics were measured. (G) Inhibited growth is now observed in batch, with more stable lysis after 8 hours.

4.3.6 Screening libraries of synthetic lysis oscillators.

To fully demonstrate the capability of this device to screen dynamic gene circuits, we generated and screened a library of synthetic oscillator strains (Figure 4.14A). We used a single-plasmid oscillator design based on the two-plasmid oscillator previously used by our lab for drug delivery [9]. The expression of the LuxI protein, which produces the quorum sensing autoinducer *N*-Acyl homoserine lactone (AHL), generates synchronized positive feedback in a colony of isogenic cells. The positive activation of the pLuxI promoter in-turn drives the expression negative

feedback *via* the lysis protein, E, causing the synchronized lysis of the colony. A library of this single-plasmid oscillator was generated by randomly mutagenizing 5 bases in the ribosome binding site (RBS) upstream of the lysis protein, E. With 1,024 possible library members, varying RBS strength will alter the translation of E, potentially changing the negative feedback, and thus, oscillatory dynamics of this system. Because of these varying oscillatory phenotypes, this oscillator is an excellent candidate for dynamic phenotype screening. Furthermore, these oscillations are only observable in continuous culture, including on the acoustic loaded device, and has both fluorescent and OD based reporting (Figure 4.14B).

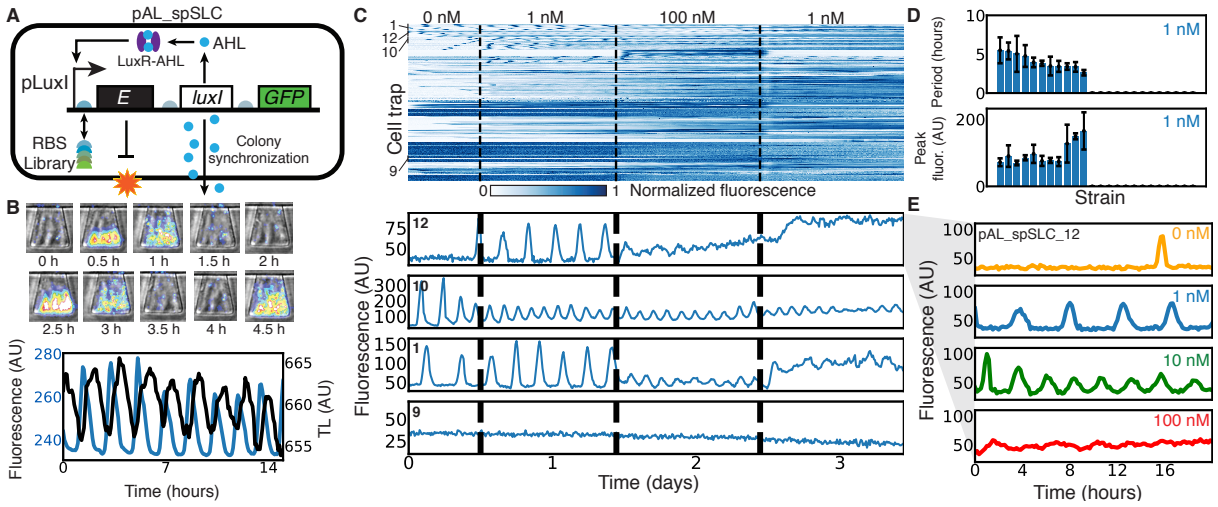

Figure 4.14: Screening of synthetic oscillator strains. (A) A single-plasmid synthetic oscillator was developed with AHL production from the LuxI protein as a cell-synchronized positive feedback mechanism, and cell lysis as negative feedback. A library was created by having randomizing four bases in the ribosome binding site (RBS) upstream of the lysis protein, E, for a total of 256 possible library members. (B) This single-plasmid oscillator performs well in the acoustic loaded microfluidic platform, with both GFP and TL used as measurable outputs. (C) 24 library members were screened on a 48 strain device and subjected to temporal changes in the background AHL concentration. Different dynamic phenotypes were measured from across these 24 strains, with four examples shown. (D) Extracted parameters of 24 oscillator strains under 1 nM AHL. The period oscillatory period of each strain is shown, with 0 indicating no measured oscillations. The mean peak fluorescence values are shown below. Error bars represent standard deviation of measured cell traps. (E) A single strain subjected to multiple background concentrations of AHL exhibits varying dynamic phenotypes.

24 randomly selected library members were run across the 48-strain DynaScreen device, with fluorescence measurements taken every 10 minutes (Figure 4.14C). Over a period of 3

days, the device was subjected to varying background concentrations of AHL to determined its effect on oscillatory dynamics. Clustering of all cell traps reveals an abundance of phenotypes, predominantly "broken" oscillators with no oscillatory dynamics, but several working oscillators. Four strains are highlighted in Figure 4.14C (sequences shown in Figure 4.15C), showing oscillators that only activate under 1 nM AHL, 100 nM AHL, or are robust to varying concentrations. Dynamic parameters (period and peak fluorescence) were extracted for all 24 strains at 1 nM AHL to demonstrate that a range of values could be generated and measured with this library and device (Figure 4.14D). Finally, Multiple chips were set up in parallel to observe the phenotype of these oscillator strains grown under different AHL concentration (Figure 4.14E). We observe that for the highlighted strain, oscillations are sparse at 0 nM of AHL, regular at 1 nM and 10 nM with frequency increasing at higher concentrations, and absent at 100 nM.

4.4 Discussion

In this Chapter, we develop a high-throughput, multiplexed microfluidic platform for continuous culture and real-time measurement of bacterial colonies. We demonstrate that with acoustic cell-loading, we can rapidly inoculate a microfluidic device from liquid culture while maintaining cell health and gene circuit phenotype. We demonstrate the parallel culture capabilities of this device for the dynamic phenotype screening of fluorescence and cell-density-based biosensors and synthetic lysis oscillators under continuous growth conditions. We use this to rapidly validate biosensor performance using both laboratory-made solutions and environmental samples dosed with heavy metals. Furthermore, we show how the dynamic screening of these gene circuits enables precise tuning of circuit output, most notably in the case of synthetic oscillators where batch-screening approaches could not be used.

Towards environmental monitoring, synthetic biologists can have the most impact harnessing the breadth of sensing targets as well as the real-time capabilities of biosensors. Well

established analytical techniques such as atomic absorption spectroscopy (AAS) or inductively coupled plasma mass spectrometry (ICP-MS) are highly sensitive, but limited by the number of detectable compounds and operate on discrete samples [86]. Whole-cell biosensors are notable for their ability to sense a wide range of compounds, previously being developed for the detection harmful compounds for environmental safety [18, 29, 59, 60, 128, 133–135], sensing metabolites relevant to bioproduction [61–64, 143], or perform medical diagnostics *in vivo* [65–67, 144]. However, while many environmental sensors are approaching real-world deployability, most are limited to batch systems [133–135] and the single measurement of a single sample or compound. Multiplexed microfluidic platforms fully utilize these properties of whole-cell biosensors, enabling the co-deployment of multiple engineered strains without the need to consider microbial ecologies [145], while operating *in situ* and detecting compounds in real time. Existing platforms have greatly advanced the capabilities on this front [29, 30], and we believe the accessibility, throughput, dynamics, and week-long time scales achievable with our device will push it even further.

Tuning genetic circuit outputs by screening variant libraries for a desired phenotype has long been fundamental to synthetic biology design. However, the mass-screening of dynamic phenotypes has remained a persistent challenge and our ability to generate genotypic diversity far exceeds our ability to screen complex phenotypes [146]. Despite limited means for dynamic phenotype screening, canonical gene circuit motifs, including oscillators, logic gates, and feedback controllers have been increasingly deployed in time-dependent applications spanning metabolic engineering to therapeutic delivery [9, 147–150]. Multiplexed microfluidics, such as ours, can aid in the development of circuits like these, for both academic research and as synthetic biology exits the lab and enters the real world. Microfluidics has served as powerful tool for approximating complex real-world environments in the past, simulating environments spanning soil to human organs [151, 152]. While not a perfect recreation of these complex environments, tuning environmental and time-dependent parameters achievable with microfluidics serves an

important role in the prototyping and scale-up of all classes of gene circuits.

4.5 Materials and methods

4.5.1 Microfluidic device development and fabrication

Our group has previously described the microfabrication techniques used to pattern SU-8 photoresist onto a silicon wafer to create the mold for our device [16]. A poly-dimethylsiloxane (PDMS) device was made from the wafer by mixing 77 grams of Sylgard 184 and pouring it on the wafer centered on a level 5"x5" glass plate surrounded with an aluminum foil seal. The degassed wafer and PDMS was cured on a flat surface for one hour at 95°C.

4.5.2 Microfluidic device loading and bonding

A PDMS device cleaned with 70% Ethanol and adhesive tape was aligned to a custom fixture compatible with the Labcyte Echo. Both the fixture and a clean glass slide sonicated with 2% Helmanex III were exposed to oxygen plasma. *E. coli* to be used in microfluidic experiments were grown for 16 hours on LB media, after which 45 μL of each strain was pipetted into a well of an Echo plate (get details) and 2.5 nL acoustically transferred to each device position on the PDMS device. The spotted device and glass slide were bonded together and cured at 37°C for two hours.

4.5.3 Microfluidic experimental protocol

Microfluidic experiments were performed on a custom optical enclosure described in the supplemental information or on a Nikon TE2000-U epifluorescent inverted microscope (Nikon Instruments Inc., Tokyo, Japan). Cells were initially grown on the device on LB media with Kanamycin, and 0.075% Tween-20 until traps were filled to confluence. The media was then

switched to an HM9 minimal media described in Table S1, which was based on a previous study [110] and optimized for microfluidic *E. coli* growth with minimal traces of metals. Cells were grown on HM9 for at least 12 hours before exposure to heavy metals. Extracted fluorescence time series were normalized to remove device background fluorescence and strain background fluorescence. Detailed methods on experimental set-up and data collection can be found in the supplement (Table S3).

4.5.4 Cloning and mutant library generation

Fluorescence-based sensor strains were cloned using Q5 DNA Polymerase (New England Biolabs, Ipswich, MA) to amplify PCR products from source plasmids, custom gBlocks (IDT, Coralville, IA), or the MG1655 *E. coli* genome and reassembled using Gibson assembly [153]. Lysis variants of these strains were cloned by amplifying the fluorescence-sensor plasmid, excluding GFP, and the E, lysis protein with 20 base pairs of homology with the amplified fluorescence-sensor vector on both the 5' and 3' ends, before joining with Gibson assembly.

Error prone PCR was performed on the promoter and RBS regions of each sensor plasmid using a previously described protocol [154], using 30 amplification cycles. Sensor strain backbones were amplified using PCR with the high-fidelity Q5 DNA Polymerase (New England Biolabs, Ipswich, MA), and assembled to mutated promoter constructs using CPEC assembly [155]. Assembled constructs were transformed into *E. coli* MG1655 and plated onto LB plates with kanamycin. 144 colonies were selected, grown on 96 well plates and subjected to metal exposures on the 144-strain microfluidic device. 48 variants with differential GFP production from wild-type were subjected to further screening on the 48-strain device. Plasmid sequences were confirmed with Sanger sequencing (Eton Bioscience, San Diego, CA). Sequences of the highlighted variants are shown in Figure 4.15.

To generate a mutant library of the pAL_spSLC oscillator plasmid, 5 base pairs in the Shine-Dalgarno sequence of the ribosome binding site (located 7 to 12 base pairs upstream of the

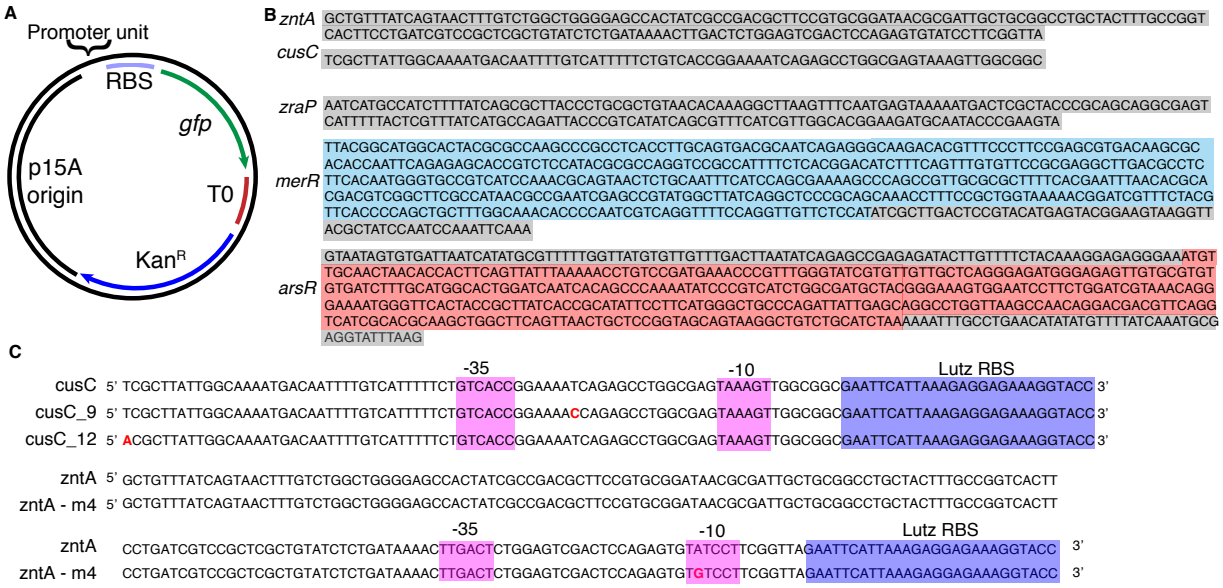

Figure 4.15: Plasmid map and sequences for biosensor strains. (A) The general plasmid map for the sensor strains used in this study. All strains have the same background, with only the promoter unit varying between each strain. (B) The different promoter units used for each of the sensor strains. The *merR* and *arsR* strains also express their own regulator, highlighted in blue and red, respectively. (C) Sequences of the promoter regions subjected to error prone PCR for the base and mutant strains. -10 and -35 sites are highlighted in pink, the RBS region is highlighted in blue, and mutated bases are highlighted in red.

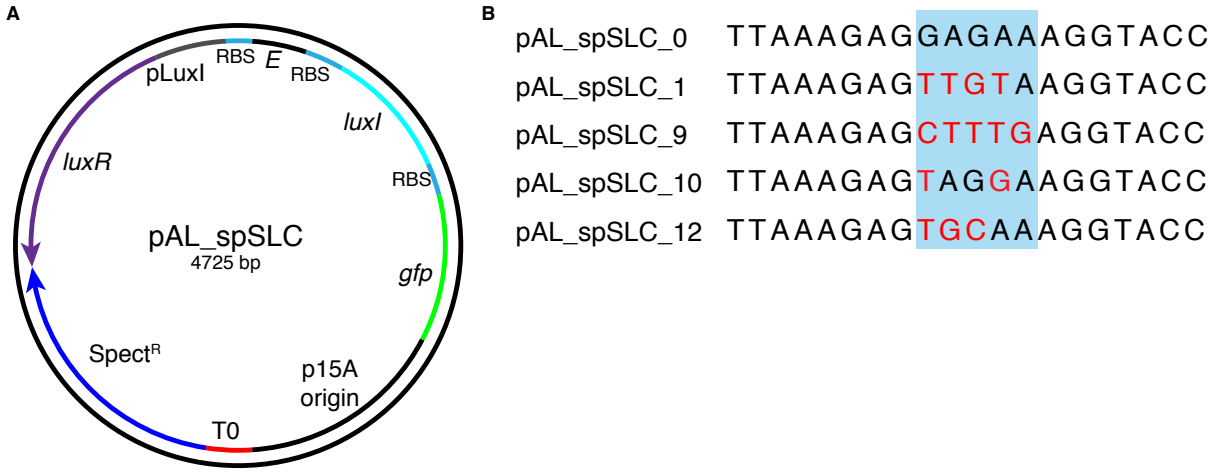

Figure 4.16: Plasmid map and sequences for single plasmid lysis oscillator circuits. (A) The general plasmid map for the single plasmid synthetic lysis circuit (spSLC) strains used in this study. (B) Sequences of the RBS region, with the 5 randomized bases highlighted in blue. Deviations from the base strain (pAL_spSLC_0) are shown in red.

start codon of the lysis protein, E) were randomized by site directed mutagenesis. The original sequence at this position was: GAGAA. First, the entire plasmid was PCR amplified with the following degenerate primers where N indicates any base: 5' CATTAAAGAGNNNNNAGGTAC-CATGATGGTAC 3' and 5' AATTCTCTCTATCACTGATAG 3'. The PCR reaction mix was incubated with DPNI at 37C for 30 minutes to digest template plasmid and then the 4.7 kb PCR product was run on an agarose gel and extracted using a QIAquick Gel Extraction Kit (QIAGEN). 1μL of the gel extracted PCR product was mixed with 0.5μL T4 ligase buffer, 0.5μL T4 PNK, and 3μL of DNase-free water and incubated at 37 °C. Next, 0.5μL T4 ligase buffer, 0.5μL T4 DNA ligase, and 4μL were added to the reaction mixture and the mixture was incubated at room temperature overnight. The following day, 50μL of chemically competent MG1655 E. coli cells were transformed with 3μL of the reaction mix and plated on an LB agar plate containing 0.2% glucose and spectinomycin. 24 colonies from the agar plate were randomly selected for mutant screening and grown up for 16 hours in LB media with 0.2% glucose and spectinomycin prior to use in experiments. Sequences of the highlighted variants are shown in Figure 4.16.

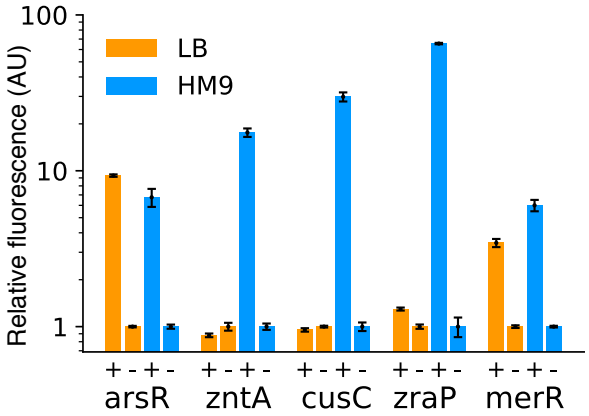

Figure 4.17: Comparison of HM9 and LB media. Plate reader analysis of the five sensor strains grown and induced with metal on either LB or HM9 growth media. Triplicate fluorescence values were measured after 16 hours of growth, and divided by the fluorescence of the uninduced counterpart. *arsR* was induced with 1.5 ppm As(III), *zntA* with 0.5 ppm Cd(II), *cusC* with 0.63 ppm Cu(II), *zraP* with 8.5 ppm Pb(II), and *merR* with 0.2 ppm pf Hg(II)

4.5.5 Media selection

HM9 was chosen for experiments because of the know characteristic of LB media to chelate metals [156]. The recipe for HM9 media is described in Table 3.3, and was based on a previous study [110] and optimized for microfluidic *E. coli* growth with minimal traces of metals. Sensor performance on this media was verified with initial inductions of the strains on both HM9 and LB. We observed that LB inhibited GFP expression as compared to HM9 for all five tested metals except As(III) (Fig. 4.17), for which there was little discernable difference. Furthermore, the low autofluorescence of HM9 reduces the background signal from media when imaging in microfluidics, making it more ideal for detecting subtle fluorescent changes in low-sensitivity strains.

4.5.6 Plate reader experiments

E. coli to be used in plate reader experiments were grown for 16 hours on LB kanamycin media, reseeded in a 1:1000 dilution into fresh media, and allowed to grow for four hours until

it reached an OD600 of about 0.1 before being added to HM9 media in a 96-well plate. For fluorescence dose-response curves, a 96-well plate was prepared with 200 μL of HM9 and a range of metal concentrations before cell culture seeding. Fluorescence and OD600 values were continuously measured during growth. The maximum OD normalized fluorescence value at the OD closest to 0.2, normalized to a promoterless strain, is presented in Figure 4.7. For lysis dose-response curves, a 96-well plate was prepared with 170 μL of HM9 before cell culture seeding. Once measured OD was roughly 0.15, 30 μL of HM9 with metal was added to bring each well to the reported metal concentration. OD values were continuously measured after metal exposure.

4.5.7 Live-cell imaging and data

Microfluidic devices were imaged in a custom optical enclosure continuously every ten minutes in both the transmitted light and gfp fluorescence channels with a 1 second and 60 second exposure respectively. The custom optical enclosure uses an SBIG STX-16803 CCD Camera with a custom lens stack assembly containing the Semrock FF01-466/40-32-D-EB and FF02-520/28-50-D-EB excitation and emission filters, respectively. The enclosure has green and blue LED spotlight sources obtained from ProPhotonix for transmitted light and fluorescence light sources, respectively. The optical resolution of the enclosure is 36 μm. The enclosure was temperature controlled to 37°C.

Images were synced from the enclosure to a server via WiFi for further data processing. Custom software produced flat-field-corrected images in both channels in real-time to remove optical vignetting using the following equation:

$$C = m * \frac{R - D}{F' - D'} \tag{4.1}$$

where R is the raw image to be flat-field corrected (Fig. 4.18), D is the dark-current image for

that device, taken at the same exposure settings as R, F' is a raw image taken by the camera with no device present, D' is the dark-current image taken at same exposure as F', m is the mean value for all values in the array $(F' - D')$, and C is the resulting corrected image.

Flat-field corrected images were then processed in ImageJ, where a custom "Region Of Interest" or ROI manager was used to extract fluorescence, transmitted light, and background values (Fig. 4.19).

Data was initially processed by subtracting the local background signal, in order to eliminate any local or regional fluctuations that are of an additive (or, analogously, subtractive) nature. The result of this background correction was to produce a vector $\vec{x}_t$ representing the background-corrected fluorescent signals of all cell traps at time t:

$$x_{(t,\, s_i)} = x^{trap}_{(t,\, s_i)} - x^{background}_{(t,\, s_i)} \tag{4.2}$$

where t refers to the current time point, s_i refers to the strain in cell trap i, $x^{trap_{(t,\, s_i)}}$ is the flat-field corrected fluorescent signal from the $trap$ of position i at time t, and $x^{background}_{(t,\, s_i)}$ refers to the flat-field corrected local background fluorescent signal at position i at time t.

Signal-to-noise ratio (SNR) was determined as the background normalized fluorescence value divided by the standard deviation of the fluorescence values between the HM9 switch and the first metal induction.

4.5.8 Source water testing

Sensor strains were tested on four different water sources: (1) water from the Alvarado Water Treatment Plant in San Diego, CA, (2) suburban tap water from San Diego, CA, (3) Ocean water from Blacks Beach in La Jolla, CA (3) and apple juice (Trader Joe's Fresh Pressed Apple Juice SKU#88463). Once a day, each device was exposed to a mixture of a single toxin dissolved in HM9 made from the respective water source for four hours. When not exposed to metal,

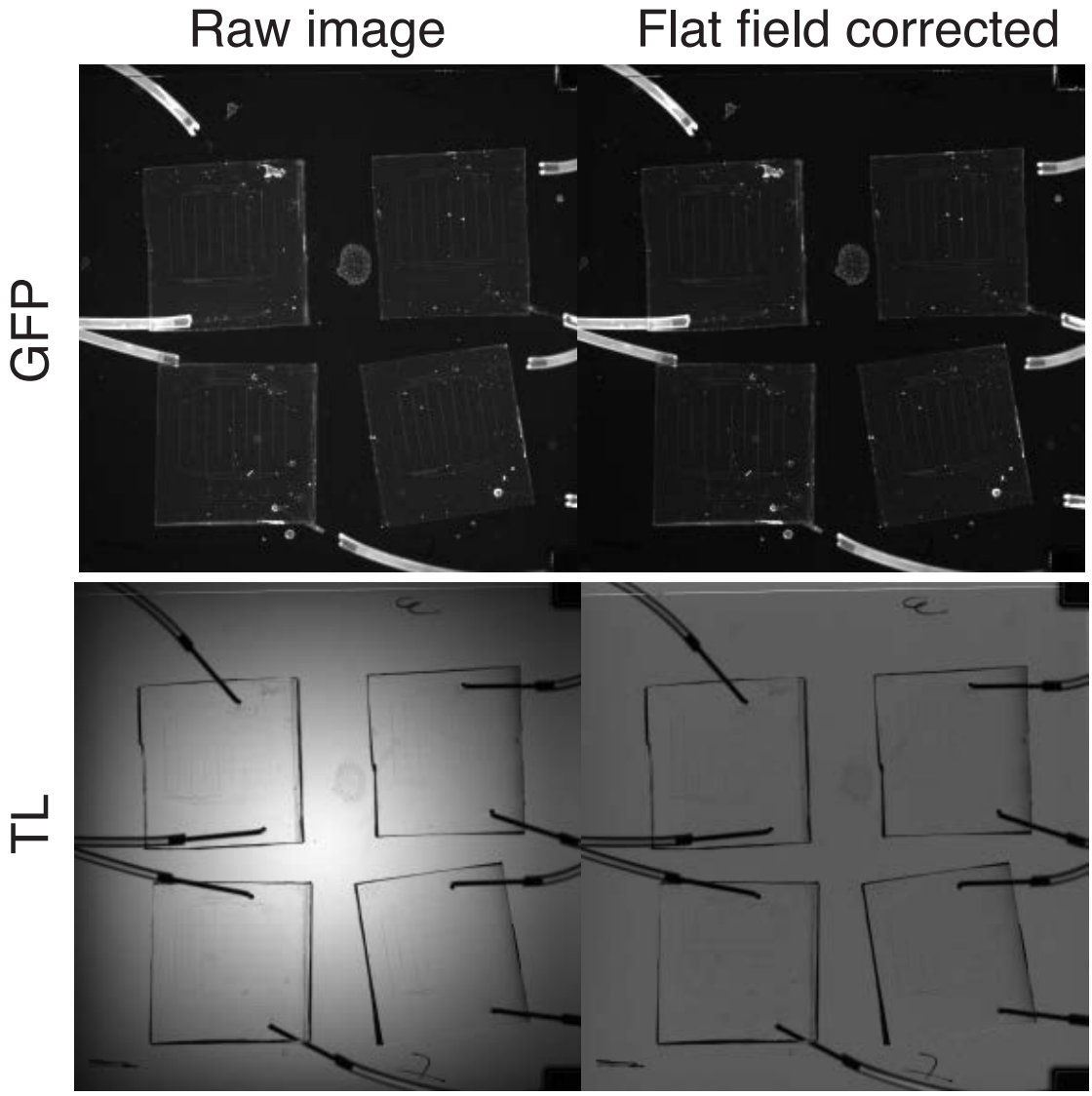

Figure 4.18: Flat field correction of fluorescence and transmitted light images taken on custom optics. The raw and flat field corrected images of four chips, for both the GFP and transmitted light (TL) channels are shown.

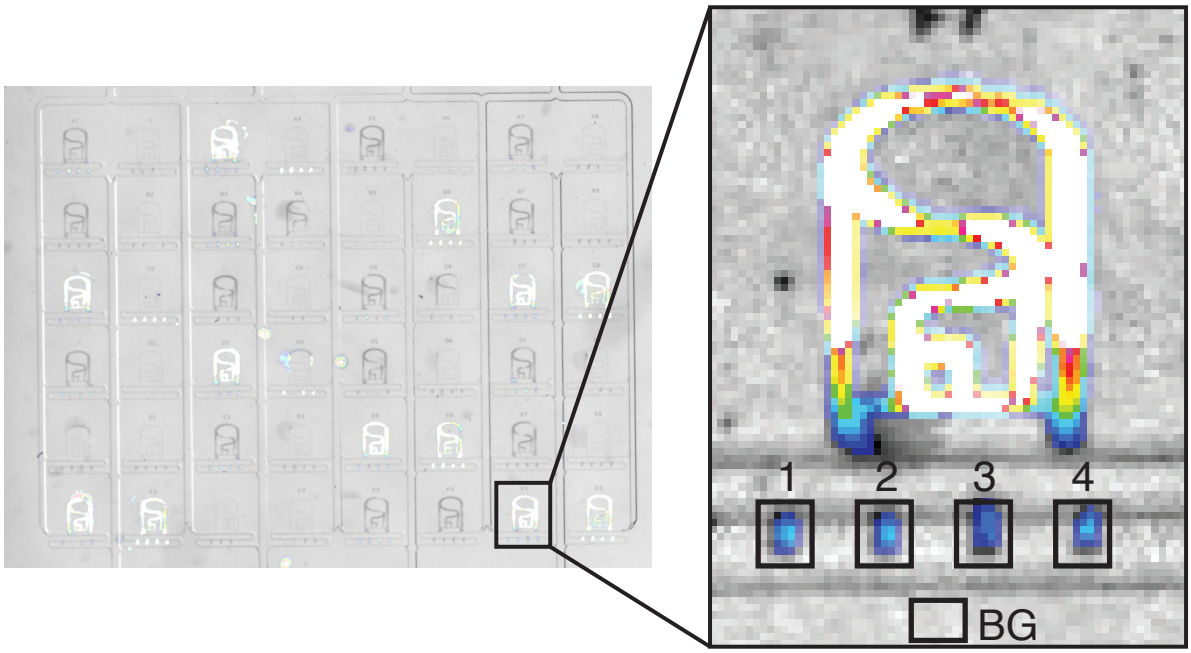

Figure 4.19: Data extraction regions from microscope images. Fluorescence and TL values are extracted from each of the four cell traps at each strain spotting position. A background value is also extracted for further normalization.

cells were grown on HM9 made from MilliQ water. Signal-to-noise ratio was determined as the fluorescence value at each time point divided by the standard deviation of fluorescence 7.5 hours before induction. Fold change was calculated from the signal-to-noise ratio of for four biological replicates of each sensor.

4.5.9 Electrochemostat measurements

The macro-scale bacterial characterization was performed using an 'electrochemostat' composed of a custom-built 3D printed housing where the electrochemical and turbidity detector-sare enclosed. This assembly was placed in a 37 °C incubator for experiments. The 3 mL bacterial culture, confined in an autoclavable beaker, is stirred at 700 rpm. Fresh media is continuously supplied with a P625/10K.143 peristaltic pump (Instech, Plymouth Meeting, PA, US) at a typical flow rate of 25 μL/min. A removable lid holds in place inlet and outlet tubings, and the disposable screen-printed electrochemical sensor consisting of a 3 mm diameter gold working electrode and a 4 mm gold counter/reference electrode (C220AT, DropSens-Metrohm, Riverview, FL, US). The lid also contains has an extra opening for aerobic conditions. A new electrode is used for each experiment and connected to a type B screen printed electrode adapter (IORodeo, Pasadena, CA, US) interfaced to the potentiostat. The optimal potential frequency for impedance measurement in these experiments was 100 MHz. Turbidity and impedance values were taken every 30 s and transmitted via USB to a PC.

4.6 Protocol for the acoustic loading of PDMS microfluidic devices with cells grown in liquid culture.

Below is the step by step loading and bonding protocol for transferring liquid cultures of *E. coli* onto microfluidic devices using a Labcyte Echo. Steps for the wafer fabrication, PDMS

fabrication, and glass slide preparation can be found in Section 2.5.

4.6.1 Cell preparation

Timing: 2-16 hours

1. Grow cells in 45 μL of media in an Echo Qualified 384-Well Polypropylene Microplate. Cells can be grown for varying periods of time ranging from an early exponential phase OD (roughly 2 hours) to a saturated overnight culture (16 hours).

2. Spin down microplate for 1 min at 100 x g.

4.6.2 Creating alignment markers

Timing: 10 minutes

1. Clean the acrylic alignment tool with 70% ethanol in water and scrub lightly with a Kimwipe or paper towel.

2. Run the Labcyte Echo Plate Reformat software for transfer 2.5-5 nL of cell culture from a 384-well source plate to a 1-well destination plate. Destination well x and y offsets will be specified manually in a custom pick list .csv.

    - Select 384PP_AQ_BP2 as the source plate type.
    - For a roughly 6.5 mm acrylic piece, specify destination plate dimensions with a plate height of 8.5 mm. Giving a space of 2 mm between acrylic piece and source plate. This distance can be increased or reduced is desired.

3. After initiating the protocol, select “Partial plates (as defined by pick list)” and import the .csv mapping file.

4. When prompted, load the 384-Well Polypropylene Microplate as the source plate, and the acrylic alignment tool as the destination plate.

4.6.3 Aligning PDMS to the acrylic tool

Timing: 5-10 minutes

1. Using a Model 200 Mask Aligner (OAI, San Jose, CA) mask aligner or similar system, place the PDMS device feature-side-down on top of the wafer chuck. Remove the tape from the non-feature side of the PDMS.

2. Place the acrylic tool on top of the mask holder, with the alignment cells facing down, above the PDMS.

3. Using the micrometers on the mask aligner, align the center of the spotting regions with the center of the alignment cells.

4. Bring the PDMS and acrylic tool into contact using the wafer chuck.

5. Remove the tape from the feature-side of the PDMS and check the alignment, ensuring that the PDMS did not shift when it came into contact with the acrylic tool. If necessary, adjust the alignment.

6. Re-tape the feature side of the PDMS until ready to expose the PDMS to oxygen plasma.

4.6.4 Oxygen plasma exposure

Timing: 10 minutes

1. Expose the clean glass slide and the PDMS acrylic stack to 30W of oxygen plasma for 30 seconds.

2. Blow any dust off the glass slide and PDMS acrylic stack with compressed nitrogen.

4.6.5 Loading and bonding the device

Timing: 10 minutes

1. Run the Labcyte Echo Plate Reformat software for transfer 2.5 nL of cell culture from a 384-well source plate to a 1-well destination plate.
    - Select 384PP_AQ_BP2 as the source plate type.
    - Specify new destination plate dimensions to account for the addition of a roughly 5.5 mm thick PDMS device. Specify destination plate dimensions with a plate height of 14 mm.
2. After initiating the protocol, select "Partial plates (as defined by pick list)" and import the same .csv mapping file used to place alignment markers.
3. When prompted, load the 384-Well Polypropylene Microplate as the source plate, and the acrylic alignment tool as the destination plate.
4. Once acoustic transfer is complete, peel the spotted PDMS off the acrylic piece and gently place it face down on the center of the oxygen-plasma-exposed glass slide.
5. Gently tap the top of the PDMS, ensuring that the device bonds to the glass.
6. Incubate the device at 37C for at least two hours.

4.7 Acknowledgements

Chapter 4 contains material being prepared for submission as A multiplexed microfluidic platform for real-time heavy metal sensing and dynamic phenotype screening. Csicsery, Nicholas, Lezia, Andrew, O'Laughlin, Richard, Stasiowski, Elizabeth, Baumgart, Leo, and Hasty, Jeff. The dissertation author was the primary author and researcher of this material.

Chapter 4 also contains material being prepared for submission as A custom high-throughput microfluidic and optical platform for the continuous culturing and measurement of *E. coli* and *S. cerevisiae*. Csicsery, Nicholas*, Stasiowski, Elizabeth*, Thouvenin, Gregoire, Graham, Garrett, O'Laughlin, Richard, Mather, William H., Ferry, Michael, Cookson, Scott, and Hasty, Jeff. (*equal contribution) The dissertation author was one of the primary authors and researchers of this material.

Chapter 4 also contains material under review at *Science Advances* as Interfacing gene circuits with microelectronics through engineered population dynamics. Din, M. Omar*, Martin, Aida*, Razinkov, Ivan, Csicsery, Nicholas, and Hasty, Jeff. (*equal contribution) The dissertation author was a secondary author and researcher of this material.

Bibliography

[1] David A Jackson, Robert H Symonst, and Paul Berg. Biochemical Method for Inserting New Genetic Information into DNA of. *Proc. Natl. Acad. Sci.*, 69(10):2904–2909, 1972.

[2] Timothy S. Gardner, Charles R. Cantor, and James J. Collins. Construction of a genetic toggle switch in Escherichia coli. *Nature*, 2000.

[3] Michael B Elowitz and Stanislas Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, pages 335–338, 2000.

[4] Drew Endy. Foundations for engineering biology. *Nature*, 438(November):449–453, 2005.

[5] Ernesto Andrianantoandro, Subhayu Basu, David K Karig, and Ron Weiss. Synthetic biology: new engineering rules for an emerging discipline. *Molecular Systems Biology*, pages 1–14, 2006.

[6] Jesse Stricker, Scott Cookson, Matthew R Bennett, William H Mather, Lev S Tsimring, and Jeff Hasty. A fast, robust and tunable synthetic gene oscillator. *Nature*, 456(7221):516–9, nov 2008.

[7] Brynne C Stanton, Alec A K Nielsen, Alvin Tamsir, Kevin Clancy, Todd Peterson, and Christopher A Voigt. Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nature Chemical Biology*, 10(2):99–105, 2013.

[8] IndustryARC. Synthetic Biology Market - Forecast (2019 - 2024). Technical report, 2019.

[9] M Omar Din, Tal Danino, Arthur Prindle, Matt Skalak, Jangir Selimkhanov, Kaitlin Allen, Ellixis Julio, Eta Atolia, Lev S Tsimring, Sangeeta N Bhatia, and Jeff Hasty. Synchronized cycles of bacterial lysis for in vivo delivery. *Nature*, 536, 2016.

[10] Felix Moser, Eléonore Tham, Lina M González, Timothy K Lu, and Christopher A Voigt. Light-Controlled , High-Resolution Patterning of Living Engineered Bacteria Onto Textiles , Ceramics , and Plastic. *Advanced Functional Materials*, 1901788:1–11, 2019.

[11] Sheeja Jagadevan, Avik Banerjee, Chiranjib Banerjee, Chandan Guria, Rameshwar Tiwari, and Mehak Baweja. Biotechnology for Biofuels Recent developments in synthetic biology

and metabolic engineering in microalgae towards biofuel production. *Biotechnology for Biofuels*, pages 1–21, 2018.

[12] Chris J Paddon and Jay D Keasling. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nature Reviews Microbiology*, (April), 2014.

[13] Ahmad S Khalil and James J Collins. Synthetic biology: applications come of age. *Nature reviews. Genetics*, 2010.

[14] Daniel Figeys and Devanand Pinto. A Revolution in Biological and Medical Sciences A look at some of the. *Analytical Chemistry*, 72(9):330–335, 2000.

[15] Daniel T. Chiu, Andrew J. DeMello, Dino Di Carlo, Patrick S. Doyle, Carl Hansen, Richard M. Maceiczyk, and Robert C.R. Wootton. Small but Perfectly Formed? Successes, Challenges, and Opportunities for Microfluidics in the Chemical and Biological Sciences. *Chem*, 2(2):201–223, 2017.

[16] M. S. Ferry, I. A. Razinkov, and J. Hasty. *Microfluidics for synthetic biology: From design to execution*, volume 497. 2011.

[17] Tal Danino, Octavio Mondragón-Palomino, Lev Tsimring, and Jeff Hasty. A synchronized quorum of genetic clocks. *Nature*, 463(7279):326–30, jan 2010.

[18] Arthur Prindle, Phillip Samayoa, Ivan Razinkov, Tal Danino, Lev S. Tsimring, and Jeff Hasty. A sensing array of radically coupled genetic 'biopixels'. *Nature*, 2012.

[19] Yang Li, Meng Jin, Richard O'Laughlin, Philip Bittihn, Lev S. Tsimring, Lorraine Pillus, Jeff Hasty, and Nan Hao. Multigenerational silencing dynamics control cell aging. *Proceedings of the National Academy of Sciences*, 114(42):11253–11258, 2017.

[20] Ivan Alexandrovich Razinkov, Bridget L Baumgartner, Matthew R Bennett, Lev S Tsimring, and Jeff Hasty. Measuring Competitive Fitness in Dynamic Environments. *The journal of physical chemistry. B*, jul 2013.

[21] Matthew R. Bennett, Wyming Lee Pang, Natalie A. Ostroff, Bridget L. Baumgartner, Sujata Nayak, Lev S. Tsimring, and Jeff Hasty. Metabolic gene regulation in a dynamically changing environment. *Nature*, 454(7208):1119–1122, 2008.

[22] J. Uhlendorf, A. Miermont, T. Delaveau, G. Charvin, F. Fages, S. Bottani, G. Batt, and P. Hersen. Long-term model predictive control of gene expression at the population and single-cell levels. *Proceedings of the National Academy of Sciences*, 109(35):14271–14276, 2012.

[23] Jerome T. Mettetal, Dale Muzzey, Gomez-Uribe Carlos, and Alexander van Oudenaarden. The Frequency Dependence of Osmo-adaptation in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci.*, 23(1):1–7, 2008.

[24] Nicolas Dénervaud, Johannes Becker, Ricard Delgado-gonzalo, Pascal Damay, Arun S Rajkumar, and Michael Unser. A chemostat array enables the spatio-temporal analysis of the yeast proteome. *Proceedings of the National Academy of Sciences*, 110(39):15842–15847, 2013.

[25] Rongfei Zhang, Haiyu Yuan, Shujing Wang, Qi Ouyang, Yong Chen, Nan Hao, and Chunxiong Luo. High-Throughput single-cell analysis for the proteomic dynamics study of the yeast osmotic stress response. *Scientific Reports*, 7(February), 2017.

[26] Won-ki Huh, James V Falvo, Luke C Gerke, Adam S Carroll, Russell W Howson, Jonathan S Weissman, and Erin K O Shea. Global analysis of protein localization in budding yeast. *Nature*, 425:686–691, 2003.

[27] Alon Zaslaver, Anat Bren, Michal Ronen, Shalev Itzkovitz, Ilya Kikoin, Seagull Shavit, Wolfram Liebermeister, Michael G. Surette, and Uri Alon. A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. *Nature Methods*, 3(8):623–628, 2006.

[28] Yuichi Taniguchi, Paul J Choi, Gene-wei Li, Huiyi Chen, Mohan Babu, Jeremy Hearn, Andrew Emili, and X Sunney Xie. Single molecule analysis of E.coli. *Science*, 329(5991):533–538, 2010.

[29] Francesca Volpetti, Ekaterina Petrova, and Sebastian J Maerkl. A Microfluidic Biodisplay. *ACS Synthetic Biology*, 6:1979–1987, 2017.

[30] Ce Zhang, Hsiung-Lin Tu, Gengjie Jia, Tanzila Mukhtar, Verdon Taylor, Andrey Rzhetsky, and Savaş Tay. Ultra-multiplexed analysis of single-cell dynamics reveals logic rules in differentiation. *Science Advances*, 5(4):eaav7959, apr 2019.

[31] Boris Kholodenko, Michael B Yaffe, and Walter Kolch. Computational approaches for analyzing information flow in biological networks. *Science signaling*, 5(220):re1, apr 2012.

[32] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network motifs: Simple building blocks of complex networks. *Science*, 2002.

[33] Galit Lahav, Nitzan Rosenfeld, Alex Sigal, Naama Geva-Zatorsky, Arnold J. Levine, Michael B. Elowitz, and Uri Alon. Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nature Genetics*, 2004.

[34] S. Hooshangi, S. Thiberge, and R. Weiss. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proceedings of the National Academy of Sciences*, 2005.

[35] Jonathan M. Raser and Erin K. O'Shea. Control of Stochasticity in Eukaryotic Gene Expression. *Science*, 304:1811–1814, 2004.

[36] Markus Krupp, Jens U. Marquardt, Ugur Sahin, Peter R. Galle, John Castle, and Andreas Teufel. RNA-Seq Atlas-a reference database for gene expression profiling in normal tissue by next-generation sequencing. *Bioinformatics*, 2012.

[37] Gioele La Manno, Ruslan Soldatov, Amit Zeisel, Emelie Braun, Hannah Hochgerner, Viktor Petukhov, Katja Lidschreiber, Maria E. Kastriti, Peter Lönnerberg, Alessandro Furlan, Jean Fan, Lars E. Borm, Zehua Liu, David van Bruggen, Jimin Guo, Xiaoling He, Roger Barker, Erik Sundström, Gonçalo Castelo-Branco, Patrick Cramer, Igor Adameyko, Sten Linnarsson, and Peter V. Kharchenko. RNA velocity of single cells, 2018.

[38] Nicholas T. Ingolia, Sina Ghaemmaghami, John R.S. Newman, and Jonathan S. Weissman. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*, 2009.

[39] Yuen Ho, Albrecht Gruhler, Adrian Heilbut, Gary D. Bader, Lynda Moore, Sally Lin Adams, Anna Millar, Paul Taylor, Keiryn Bennett, Kelly Boutilier, Lingyun Yang, Cheryl Wolting, Ian Donaldson, Søren Schandorff, Juanita Shewnarane, Mai Vo, Joanne Taggart, Marilyn Goudreault, Brenda Muskat, Cris Alfarano, Danielle Dewar, Zhen Lin, Katerina Michalickova, Andrew R. Willems, Holly Sassi, Peter A. Nielsen, Karina J. Rasmussen, Jens R. Andersen, Lene E. Johansen, Lykke H. Hansen, Hans Jespersen, Alexandre Podtelejnikov, Eva Nielsen, Janne Crawford, Vibeke Poulsen, Birgitte D. Sørensen, Jesper Matthiesen, Ronald C. Hendrickson, Frank Gleeson, Tony Pawson, Michael F. Moran, Daniel Durocher, Matthias Mann, Christopher W.V. Hogue, Daniel Figeys, and Mike Tyers. Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry, 2002.

[40] D A Lashkari, J L DeRisi, J H McCusker, A F Namath, C Gentile, S Y Hwang, P O Brown, and R W Davis. Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc. Natl. Acad. Sci.*, 1997.

[41] Michael J. Heller. DNA Microarray Technology: Devices, Systems, and Applications. *Annu Rev Biomed Eng.*, 4:129–153, 2002.

[42] David L Shis, Matthew R Bennett, and Oleg A Igoshin. Dynamics of Bacterial Gene Regulatory Networks. *Annu. Rev. Biophys*, 4721(21):1–21, 2018.

[43] Teresa M. Przytycka, Mona Singh, and Donna K. Slonim. Toward the dynamic interactome: It's about time. *Briefings in Bioinformatics*, 11(1):15–29, 2010.

[44] Roger P. Alexander, Philip M. Kim, Thierry Emonet, and Mark B. Gerstein. Understanding modularity in molecular networks requires dynamics. *Science Signaling*, 2(81), 2009.

[45] Gurkan Bebek, Mehmet Koyutürk, Nathan D. Price, and Mark R. Chance. Network biology methods integrating biological data for translational science. *Briefings in Bioinformatics*, 2012.

[46] Maria Secrier and Reinhard Schneider. Visualizing time-related data in biology, a review. *Briefings in Bioinformatics*, 2013.

[47] Yolanda T. Chong, Judice L. Y. Koh, Helena Friesen, Supipi Kaluarachchi Duffy, Michael J. Cox, Alan Moses, Jason Moffat, Charles Boone, and Brenda J. Andrews. Yeast Proteome Dynamics from Single Cell Imaging and Automated Analysis. *Cell*, 161:1413–1424, 2015.

[48] Jeremy E. Purvis and Galit Lahav. Encoding and decoding cellular information through signaling dynamics. *Cell*, 152(5):945–956, 2013.

[49] Todd Thorsen, Sebastian J. Maerkl, and Stephen R. Quake. Microfluidic large-scale integration. *Science*, 298(5593):580–584, 2002.

[50] Rafael Gomez-Sjoberg, Anne A Leyrat, Dana M Pirone, Christopher S Chen, and Stephen R Quake. Versatile , Fully Automated , Microfluidic Cell Culture System. *Analytical Chemistry*, 79(22):8557–8563, 2007.

[51] Wei Sha, Ana M. Martins, Reinhard Laubenbacher, Pedro Mendes, and Vladimir Shulaev. The Genome-Wide Early Temporal Response of Saccharomyces cerevisiae to Oxidative Stress Induced by Cumene Hydroperoxide. *PLoS ONE*, 8(9):1–15, 2013.

[52] Sina Ghaemmaghami, Won Ki Huh, Kiowa Bower, Russell W. Howson, Archana Belle, Noah Dephoure, Erin K. O'Shea, and Jonathan S. Weissman. Global analysis of protein expression in yeast. *Nature*, 425(6959):737–741, 2003.

[53] Gordon J. Bean, Philipp A. Jaeger, Sondra Bahr, and Trey Ideker. Development of ultra-high-density screening tools for microbial "omics". *PLoS ONE*, 9(1), 2014.

[54] Tong Zhang and Herbert H.P. Fang. Quantification of Saccharomyces cerevisiae viability using BacLight. *Biotechnology letters*, 26(12):989–992, 2004.

[55] Reiko Hirasawa and Kumio Yokoigawa. Leavening ability of baker's yeast exposed to hyperosmotic media. *FEMS Microbiology Letters*, 194(2):159–162, 2001.

[56] Charlemagne Gilles Hounsa, E. Vincent Brandt, Johan Thevelein, Stefan Hohmann, and Bernard A. Prior. Role of trehalose in survival of Saccharomyces cerevisiae under osmotic stress. *Microbiology*, 144(3):671–680, 1998.

[57] Matthieu Jules, Gemma Beltran, Jean François, and Jean Luc Parrou. New insights into trehalose metabolism by Saccharomyces cerevisiae: NTH2 encodes a functional cytosolic trehalase, and deletion of TPS1 reveals Ath1p-dependent trehalose mobilization. *Applied and Environmental Microbiology*, 74(3):605–614, 2008.

[58] Audrey P. Gasch, Paul T. Spellman, Camilla M. Kao, Orna Carmel-Harel, Michael B. Eisen, Gisela Storz, David Botstein, and Patrick O Brown. Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes. *Molecular Biology of the Cell*, 32(11):16–20, 2002.

[59] Hyun Ju Kim, Haeyoung Jeong, and Sang Jun Lee. Synthetic biology for microbial heavy metal biosensors. *Analytical and Bioanalytical Chemistry*, 410:1191–1203, 2018.

[60] Davide Merulla and Jan Roelof Van Der Meer. Regulatable and Modulable Background Expression Control in Prokaryotic Synthetic Circuits by Auxiliary Repressor Binding Sites. *ACS Synthetic Biology*, 5(1):36–45, 2016.

[61] Jameson K. Rogers and George M. Church. Genetically encoded sensors enable real-time observation of metabolite production. *Proceedings of the National Academy of Sciences*, 113(9):2388–2393, 2016.

[62] Sambandam Ravikumar, Mary Grace Baylon, Si Jae Park, and Jong il Choi. Engineered microbial biosensors based on bacterial two-component systems as synthetic biotechnology platforms in bioremediation and biorefinery. *Microbial Cell Factories*, 16(1):1–10, 2017.

[63] Irisappan Ganesh, Sambandam Ravikumar, Seung Hwan Lee, Si Jae Park, and Soon Ho Hong. Engineered fumarate sensing Escherichia coli based on novel chimeric two-component system. *Journal of Biotechnology*, 168(4):560–566, 2013.

[64] Liangpo Li, Ran Tu, Guotian Song, Jie Cheng, Wujiu Chen, Lin Li, Lixian Wang, and Qinhong Wang. Development of a Synthetic 3-Dehydroshikimate Biosensor in Escherichia coli for Metabolite Monitoring and Genetic Screening. *ACS synthetic biology*, 8(2):297–306, 2019.

[65] Jonathan W. Kotula, S. Jordan Kerns, Lev A. Shaket, Layla Siraj, James J. Collins, Jeffrey C. Way, and Pamela A. Silver. Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proceedings of the National Academy of Sciences*, 111(13):4838–4843, 2014.

[66] Tal Danino, Kaitlin Allen, Matthew Skalak, Sangeeta N Bhatia, Arthur Prindle, Howard Li, Jeff Hasty, and Gabriel A Kwong. Programmable probiotics for detection of cancer in urine. *Science Translational Medicine*, 7(289):289ra84–289ra84, 2015.

[67] Carlos Piñero-Lambea, Gustavo Bodelón, Rodrigo Fernández-Periáñez, Angel M. Cuesta, Luis Álvarez-Vallina, and Luis Ángel Fernández. Programming controlled adhesion of E. coli to target surfaces, cells, and tumors with synthetic adhesins. *ACS Synthetic Biology*, 4(4):463–473, 2015.

[68] James A Imlay. The mismetallation of enzymes during oxidative stress. *The Journal of biological chemistry*, 289(41):28121–8, oct 2014.

[69] Pete Chandrangsu, Christopher Rensing, and John D. Helmann. Metal homeostasis and resistance in bacteria. *Nature Reviews Microbiology*, 15(6):338–350, mar 2017.

[70] Kanwal Rehman, Fiza Fatima, Iqra Waheed, and Muhammad Sajid Hamid Akash. Prevalence of exposure of heavy metals and their impact on health consequences. *Journal of Cellular Biochemistry*, 119(1):157–184, 2018.

[71] K. Jomova, Z. Jenisova, M. Feszterova, S. Baros, J. Liska, D. Hudecova, C. J. Rhodes, and M. Valko. Arsenic: Toxicity, oxidative stress and human disease. *Journal of Applied Toxicology*, 31(2):95–107, 2011.

[72] Johannes Godt, Franziska Scheidig, Christian Grosse-Siestrup, Vera Esche, Paul Brandenburg, Andrea Reich, and David A. Groneberg. The toxicity of cadmium and resulting hazards for human health. *Journal of Occupational Medicine and Toxicology*, 1(1):1–6, 2006.

[73] Robin A. Bernhoft. Mercury toxicity and treatment: A review of the literature. *Journal of Environmental and Public Health*, 2012, 2012.

[74] Max Costa and Catherine B. Klein. Toxicity and carcinogenicity of chromium compounds in humans. *Critical Reviews in Toxicology*, 36(2):155–163, 2006.

[75] D A Gidlow. Lead toxicity. *Occupational Medicine*, 54(2):76–81, 2004.

[76] United States Environmental Protection Agency. Basic Information about Lead in Drinking Water, 2017.

[77] J. Nriagu. Zinc Toxicity in Humans. *Encyclopedia of Environmental Health*, (c):801–807, 2007.

[78] Antonello Pietrangelo. Iron and the liver. *Liver International*, 36(November 2015):116–123, 2016.

[79] Lisa M. Gaetke and Ching Kuang Chow. Copper toxicity, oxidative stress, and antioxidant nutrients. *Toxicology*, 189(1-2):147–163, 2003.

[80] Mona Hanna-Attisha, Jenny LaChance, Richard Casey Sadler, and Allison Champney Schnepp. Elevated blood lead levels in children associated with the flint drinking water crisis: A spatial analysis of risk and public health response. *American Journal of Public Health*, 106(2):283–290, 2016.

[81] Dustin Renwick. Five years on, the Flint water crisis is nowhere near over. *National Geographic*, 2019.

[82] M.B. Pell and J. Schneyer. The thousands of U . S . locales where lead poisoning is worse than in Flint, 2016.

[83] Karletta Chief, Janick F Artiola, Paloma Beamer, Sarah T Wilkinson, and Raina M Maier. Understanding the Gold King Mine Spill. *Superfund Research Program*, (August):1–7, 2016.

[84] Alysa Landry. Navajo Nation Seeks $ 160 Million in Damages for Gold King Mine Spill, 2016.

[85] Luca Pujol, David Evrard, Karine Groenen-Serrano, Mathilde Freyssinier, Audrey Ruffien-Cizsak, and Pierre Gros. Electrochemical sensors and devices for heavy metals assay in water: The French groups' contribution. *Frontiers in Chemistry*, 2(APR):1–24, 2014.

[86] Asli Baysal, Nil Ozbek, and Suleyman Akm. Determination of Trace Metals in Waste Water and Their Removal Processes. *Waste Water - Treatment Technologies and Recent Analytical Developments*, 2013.

[87] Zachary D. Stephens, Skylar Y. Lee, Faraz Faghri, Roy H. Campbell, Chengxiang Zhai, Miles J. Efron, Ravishankar Iyer, Michael C. Schatz, Saurabh Sinha, and Gene E. Robinson. Big data: Astronomical or genomical? *PLoS Biology*, 13(7):1–11, 2015.

[88] Jianzhu Ma, Michael Ku Yu, Samson Fong, Keiichiro Ono, Eric Sage, Barry Demchak, Roded Sharan, and Trey Ideker. Using deep learning to model the hierarchical structure and function of a cell. *Nature Methods*, 15(4):290–298, 2018.

[89] Jason H. Yang, Sarah N. Wright, Meagan Hamblin, Douglas McCloskey, Miguel A. Alcantar, Lars Schrübbers, Allison J. Lopatkin, Sangeeta Satish, Amir Nili, Bernhard O. Palsson, Graham C. Walker, and James J. Collins. A White-Box Machine Learning Approach for Revealing Antibiotic Mechanisms of Action. *Cell*, pages 1649–1661, 2019.

[90] Jian Zhou and Olga G Troyanskaya. Predicting effects of noncoding variants with deep learning-based sequence model. *Nature methods*, 12(10):931–4, 2015.

[91] Scott Lundberg and Su-In Lee. A Unified Approach to Interpreting Model Predictions. *ArXiv e-prints*, (Section 2):1–10, 2017.

[92] Scott M. Lundberg, Gabriel G. Erion, and Su-In Lee. Consistent Individualized Feature Attribution for Tree Ensembles. *ArXiv e-prints*, (2), 2018.

[93] Rakesh Sharma, Christopher Rensing, P Rosen, Bharati Mitra, and Barry P Rosen. The ATP Hydrolytic Activity of Purified ZntA , a Pb(II)/Cd(II)/Zn(II)-translocating ATPase from Escherichia coli. *The journal of biological chemistry*, 275(Ii):3873–3878, 2000.

[94] G. Grass and C. Rensing. CueO is a multi-copper oxidase that confers copper tolerance in Escherichia coli. *Biochemical and Biophysical Research Communications*, 286(5):902–908, 2001.

[95] Sankha Ghatak, Zachary A. King, Anand Sastry, and Bernhard O. Palsson. The y-ome defines the 35% of Escherichia coli genes that lack experimental evidence of function. *Nucleic acids research*, 47(5):2446–2454, 2019.

[96] Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. *ArXiv e-prints*, 2016.

[97] Sepp Hochreiter and Jurgen Schmidhuber. Long Short-Term Memory. *Neural Computation*, 9(8):1–32, 1997.

[98] Jonathan P. McHugh, Francisco Rodríguez-Quiñones, Hossein Abdul-Tehrani, Dimitri A. Svistunenko, Robert K. Poole, Chris E. Cooper, and Simon C. Andrews. Global Iron-dependent Gene Regulation in Escherichia coli. *Journal of Biological Chemistry*, 278(32):29478–29486, 2003.

[99] Christopher Rensing, Bharati Mitra, and Barry P Rosen. The zntA gene of Escherichia coli encodes a Zn(II)-translocating P-type ATPase. *Biochemistry*, 94(December):14326–14331, 1997.

[100] Scott M Lundberg, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee. Explainable AI for Trees: From Local Explanations to Global Understanding. *arXiv preprint arXiv:1905.04610*, 2019.

[101] Samuel Dodge and Lina Karam. A study and comparison of human and deep learning recognition performance under visual distortions. *2017 26th International Conference on Computer Communications and Networks, ICCCN 2017*, 2017.

[102] Paul B Tchounwou, Clement G Yedjou, Anita K Patlolla, and Dwayne J Sutton. Molecular, clinical and environmental toxicology: v.2: Clinical toxicology. *Choice Reviews Online*, 47(10):47–5683–47–5683, 2013.

[103] Maura Allaire, Haowei Wu, and Upmanu Lall. National trends in drinking water quality violations. *Proceedings of the National Academy of Sciences*, 115(9):2078–2083, 2018.

[104] Catherine Lozupone and Rob Knight. UniFrac: A new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, 71(12):8228–8235, 2005.

[105] Frans J. De Bruijn. *Stress and Environmental Regulation of Gene Expression and Adaptation in Bacteria*. 2016.

[106] Anyou Wang and David E Crowley. Global gene expression responses to cadmium toxicity in Escherichia coli. *Journal of bacteriology*, 187(9):3259–66, may 2005.

[107] Stephen P. LaVoie and Anne O. Summers. Transcriptional responses of Escherichia coli during recovery from inorganic or organic mercury exposure. *BMC Genomics*, 2018.

[108] K. R. Brocklehurst and A. P. Morby. Metal-ion tolerance in Escherichia coli: Analysis of transcriptional profiles by gene-array technology. *Microbiology*, 2000.

[109] Leslie A. Pratt and Roberto Kolter. Genetic analysis of Escherichia coli biofilm formation: Roles of flagella, motility, chemotaxis and type I pili. *Molecular Microbiology*, 30(2):285–293, 1998.

[110] Robert A. LaRossa, Dana R. Smulski, and Tina K. Van Dyk. Interaction of lead nitrate and cadmium chloride with Escherichia coli K-12 and Salmonella typhimurium global regulatory mutants. *Journal of Industrial Microbiology*, 14(3-4):252–258, 1995.

[111] Huaiyu Mi, Anushya Muruganujan, Dustin Ebert, Xiaosong Huang, and Paul D Thomas. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic acids research*, 47(D1):D419–D426, jan 2019.

[112] Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando De Freitas. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2016.

[113] Zachary Chase Lipton, Charles Elkan, and Balakrishnan Narayanaswamy. Thresholding Classifiers to Maximize F1 Score. *ArXiv e-prints*, 2014.

[114] Alon Zaslaver, Avi E. Mayo, Revital Rosenberg, Pnina Bashkin, Hila Sberro, Miri Tsalyuk, Michael G. Surette, and Uri Alon. Just-in-time transcription program in metabolic pathways. *Nature Genetics*, 2004.

[115] Laura S. Busenlehner, Mario A. Pennella, and David P. Giedroc. The SmtB/ArsR family of metalloregulatory transcriptional repressors: Structural insights into prokaryotic metal resistance. *FEMS Microbiology Reviews*, 27(2-3):131–143, 2003.

[116] Tamar Barkay, Susan M. Miller, and Anne O. Summers. Bacterial mercury resistance from atoms to ecosystems. *FEMS Microbiology Reviews*, 27(2-3):355–384, 2003.

[117] Joe J. Harrison, Howard Ceri, and Raymond J. Turner. Multimetal resistance and tolerance in microbial biofilms. *Nature Reviews Microbiology*, 5(12):928–938, 2007.

[118] Xiguang Chen, Rowena T. Romano, and Ruihong Zhang. Anaerobic digestion of food wastes for biogas production. *International Journal of Agricultural and Biological Engineering*, 3(4):61–72, 2010.

[119] Lucas Tadeu Fuess, Antonio Djalma Nunes Ferraz Junior, Carla Botelho Machado, and Marcelo Zaiat. Temporal dynamics and metabolic correlation between lactate-producing and hydrogen-producing bacteria in sugarcane vinasse dark fermentation : The key role of lactate. *Bioresource Technology*, 247:426–433, 2018.

[120] Isabelle Dufau, Céline Frongia, Flavie Sicard, Laure Dedieu, Pierre Cordelier, Frédéric Ausseil, Bernard Ducommun, and Annie Valette. Multicellular tumor spheroid model to evaluate spatio-temporal dynamics effect of chemotherapeutics: Application to the gemcitabine/CHK1 inhibitor combination in pancreatic cancer. *BMC Cancer*, 12:1–11, 2012.

[121] Julia Oh, Allyson L. Byrd, Morgan Park, Heidi H. Kong, and Julia A. Segre. Temporal Stability of the Human Skin Microbiome. *Cell*, 165(4):854–866, 2016.

[122] Elizabeth L. Rieke, Michelle L. Soupir, Thomas B. Moorman, Fan Yang, and Adina C. Howe. Temporal Dynamics of Bacterial Communities in Soil and Leachate Water After Swine Manure Application. *Frontiers in Microbiology*, 9(December):1–11, 2018.

[123] Nan Hao, Bogdan A Budnik, Jeremy Gunawardena, and Erin K. O'Shea. Tunable signal processing through modular control of transcription factor translocation. *Science*, (339):460–464, 2013.

[124] William P Janzen. *High Throughput Screening*. Springer Protocols, 2016.

[125] Spencer R. Scott and Jeff Hasty. Quorum Sensing Communication Modules for Microbial Consortia. *ACS Synthetic Biology*, 5(9):969–977, 2016.

[126] Philipp A. Jaeger, Lilia Ornelas, Cameron McElfresh, Lily R. Wong, Randolph Y. Hampton, and Trey Ideker. Systematic Gene-to-Phenotype Arrays: A High-Throughput Technique for Molecular Phenotyping. *Molecular Cell*, 69(2):321–333.e3, 2018.

[127] John Joslin, James Gilligan, Paul Anderson, Catherine Garcia, Orzala Sharif, Janice Hampton, Steven Cohen, Miranda King, Bin Zhou, Shumei Jiang, Christopher Trussell, Robert Dunn, John W. Fathman, Jennifer L. Snead, Anthony E. Boitano, Tommy Nguyen, Michael Conner, Mike Cooke, Jennifer Harris, Ed Ainscow, Yingyao Zhou, Chris Shaw, Dan Sipes, James Mainquist, and Scott Lesley. A Fully Automated High-Throughput Flow Cytometry Screening System Enabling Phenotypic Drug Discovery. *SLAS Discovery*, 23(7):697–707, 2018.

[128] Xinyi Wan, Francesca Volpetti, Ekaterina Petrova, Chris French, Sebastian J Maerkl, and Baojun Wang. Cascaded amplifying circuits enable ultrasensitive cellular sensors for toxic metals. *Nature Chemical Biology*, 15(May), 2019.

[129] Jina Yang, Beomhee Kim, Gi Yeon Kim, Gyoo Yeol Jung, and Sang Woo Seo. Synthetic biology for evolutionary engineering: From perturbation of genotype to acquisition of desired phenotype. *Biotechnology for Biofuels*, 12(1):1–14, 2019.

[130] Daniela Quaglia, Maximilian C.C.J.C. Ebert, Paul F. Mugford, and Joelle N. Pelletier. Enzyme engineering: A synthetic biology approach for more effective library generation and automated high-throughput screening. *PLoS ONE*, 12(2):1–14, 2017.

[131] Chuankai Cheng, Edward J. O'Brien, Douglas McCloskey, Jose Utrilla, Connor Olson, Ryan A. LaCroix, Troy E. Sandberg, Adam M. Feist, Bernhard O. Palsson, and Zachary A. King. Laboratory evolution reveals a two-dimensional rate-yield tradeoff in microbial metabolism. *PLoS computational biology*, 15(6):e1007066, 2019.

[132] N. Miranda Nebane, Tatjana Coric, Sara McKellip, LaKeisha Woods, Melinda Sosa, Lynn Rasmussen, Mary-Ann Bjornsti, and E. Lucille White. Acoustic Droplet Ejection Technology and Its Application in High-Throughput RNA Interference Screening. *Journal of Laboratory Automation*, 21(1):198–203, 2016.

[133] Shimshon Belkin, Sharon Yagur-kroll, Yossef Kabessa, Victor Korouma, Tali Septon, Yonatan Anati, Cheinat Zohar-perez, Zahi Rabinovitz, Amos Nussinovitch, and Aharon J Agranat. Remote detection of buried landmines using a bacterial sensor. *Nature Biotechnology*, 35(4):2–4, 2017.

[134] Vivian Hsiu-chuan Liao, Ming-te Chien, Yuen-yi Tseng, and Kun-lin Ou. Assessment of heavy metal bioavailability in contaminated sediments and soils using green fluorescent protein-based bacterial biosensors. *Environmental Pollution*, 142, 2006.

[135] Nina Buffi, Davide Merulla, Julien Beutier, Fanny Barbaud, Siham Beggah, Harald Van Lintel, Philippe Renaud, Jan Roelof, and Van Der Meer. Lab on a Chip Development of a microfluidics biosensor for agarose-bead immobilized Escherichia coli bioreporter cells for arsenite detection in aqueous samples †. *Lab on a Chip*, pages 2369–2377, 2011.

[136] Caroline Diorio, J I E Cai, J O Y Marmor, Rowen Shinder, and Michael S D U Bow. 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, 1995.

[137] Soon-jung Park, J O Y Wireman, and Anne Summers. Genetic Analysis of the Tn2l Operator-Promoter. *Journal of bacteriology*, 174(7):2160–2171, 1992.

[138] Christopher Rensing, Bharati Mitra, and Barry P Rosen. The zntA gene of Escherichia coli encodes a Zn(II)-translocating P-type ATPase. *Biochemistry*, 94(December):14326–14331, 1997.

[139] George P Munson, Deborah L Lam, and F Wayne Outten. Identification of a Copper-Responsive Two-Component System on the Chromosome of Escherichia coli K-12. *Journal of bacteriology*, 182(20):5864–5871, 2000.

[140] Susanne Leonhartsberger, Andrea Huber, Friedrich Lottspeich, and August Bo. The hydH / G Genes from Escherichia coli Code for a Zinc and Lead Responsive Two-component Regulatory System. *Journal of Molecular Biology*, 4:93–105, 2001.

[141] Kevin D Young and R Y Young. Lytic Action of Cloned X174 Gene E. *Journal of Virology*, 44(3):993–1002, 1982.

[142] Anke Wackwitz, Hauke Harms, Antonis Chatzinotas, Uta Breuer, Christelle Vogne, and Jan Roelof Van Der Meer. Internal arsenite bioassay calibration using multiple bioreporter cell lines. *Microbial Biotechnology*, 1(2):149–157, 2008.

[143] Cláudio R. Frazão, Victor Maton, Jean M. François, and Thomas Walther. Development of a Metabolite Sensor for High-Throughput Detection of Aldehydes in Escherichia Coli. *Frontiers in Bioengineering and Biotechnology*, 6:118, aug 2018.

[144] Shimyn Slomovic, Keith Pardee, and James J Collins. Synthetic biology devices for in vitro and in vivo diagnostics. *Proc. Natl. Acad. Sci.*, 112(47):14429–14435, 2015.

[145] Nicholas S. McCarty and Rodrigo Ledesma-Amaro. Synthetic Biology Tools to Engineer Microbial Communities for Biotechnology. *Trends in Biotechnology*, 37(2):181–197, 2019.

[146] Yolanda Schaerli and Mark Isalan. Building synthetic gene circuits from combinatorial libraries: Screening and selection strategies. *Molecular BioSystems*, 9(7):1559–1567, 2013.

[147] Stephanie J. Doong, Apoorv Gupta, and Kristala L.J. Prather. Layered dynamic regulation for improving metabolic pathway productivity in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 115(12):2964–2969, 2018.

[148] Stephanie K. Aoki, Gabriele Lillacci, Ankit Gupta, Armin Baumschlager, David Schweingruber, and Mustafa Khammash. A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature*, 570(7762):533–537, 2019.

[149] Arturo Urrios, Eva Gonzalez-Flo, David Canadell, Eulàlia De Nadal, Javier Macia, and Francesc Posas. Plug-and-Play Multicellular Circuits with Time-Dependent Dynamic Responses. *ACS Synthetic Biology*, 7(4):1095–1104, 2018.

[150] Felix Moser, Amin Espah Borujeni, Amar N. Ghodasara, Ewen Cameron, Yongjin Park, and Christopher A. Voigt. Dynamic control of endogenous metabolism with combinatorial logic circuits. *Molecular Systems Biology*, 14(11):1–18, 2018.

[151] Boyang Zhang, Anastasia Korolj, Benjamin Fook Lun Lai, and Milica Radisic. Advances in organ-on-a-chip engineering. *Nature Reviews Materials*, 3(8):257–278, 2018.

[152] Roberto Rusconi, Melissa Garren, and Roman Stocker. Microfluidics Expanding the Frontiers of Microbial Ecology. *Annu. Rev. Biophys*, pages 65–91, 2014.

[153] Daniel G. Gibson, Lei Young, Ray Yuan Chuang, J. Craig Venter, Clyde A. Hutchison, and Hamilton O. Smith. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6(5):343–345, 2009.

[154] David S. Wilson and David S. Keefe. Random Mutagenesis by PCR. *Current Protocols in Molecular Biology*, pages 8.3.1–8.3.9, 2001.

[155] Jiayuan Quan and Jingdong Tian. Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS ONE*, 4(7), 2009.

[156] Sean C. Booth, Iain F.S. George, Davide Zannoni, Martina Cappelletti, Gavin E. Duggan, Howard Ceri, and Raymond J. Turner. Effect of aluminium and copper on biofilm development of Pseudomonas pseudoalcaligenes KF707 and P. fluorescens as a function of different media compositions. *Metallomics*, 5(6):723–735, 2013.