Programming Synthetic Microbial Communities for Coexistence, Coordination, and Information Processing

Adam G. Krieger

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Cellular and Molecular Biology)
in the University of Michigan
2020

Doctoral Committee:

Associate Professor Xiaoxia Nina Lin, Chair
Professor Bradley Cardinale
Professor Michael Imperiale
Associate Professor Allen Liu
Assistant Professor Yan Zhang

Adam G. Krieger

agkrieg@umich.edu

ORCID iD: 0000-0002-2891-0829

Acknowledgements

This work has only been made possible by the support and guidance of many people. I would like to thank my advisor Prof. Xiaoxia Nina Lin and all the members of the Lin lab, past and present, for excellent scientific discussions and ideas. I would like to thank the directors, administrators, fellows, associates, and co-mentor Prof. Brad Cardinale of the Integrated Training in Microbial Systems fellowship for welcoming me into their program, being so generous with their time and resources, and always being excited about studying microbial communities. I would also like to thank the directors, associate directors, affiliated professors, administrators, and students of the Cellular and Molecular Biology program for their support, both financial and moral. I would sincerely like to thank the members of my doctoral committee for providing excellent guidance and input with respect to my dissertation work and career development. I would also like to thank the members of the MCDB Microbial "Supergroup" for allowing me to attend their meetings and providing valuable input on my projects and presentations. I would like to thank Profs. John Little and David Friedman for invaluable advice regarding bacteriophage lambda. I would also like the thank my undergraduate research mentors, Profs. Ursula Jakob and Rich Hume, for sparking my interest in research and encouraging my initial efforts. Finally, I need to thank my parents, Roger and Barbara Krieger, for supporting and loving me unconditionally through many trials and tribulations. The lessons I have learned from them throughout my life about everything from

life to science have been integral to the conceptualization and implementation of this work. I need to thank my brother Brian Krieger for supporting, loving, and inspiring me. And I need to thank my incredible group of friends, John, Mikey, Mark, Mike, Eric, Amber, Emily, Anna, Shelley, Krystina, L'Meese, Phil, Brandon, Rob, Tommy, Dane, Trevor, Seany, Justin, Ben, Scott, Steven, Allie, Mickie, Alison, and Amy for supporting and loving me through a number of very different stages of my life. Thank you all.

Table of Contents

List of Figures

List of Tables

List of Appendices

Abstract

Synthetic microbial communities offer a variety of potential advantages over single species approaches for many medical, industrial, and environmental applications. At the cellular level, metabolic pathways can be distributed amongst several community residents to lower the metabolic burden on individual cells and to enable optimization of reaction conditions for different parts of metabolic pathways. At the population level, diverse microbial communities in different natural contexts have been shown to be more productive, efficient, stable, and resistant to invasion by foreign agents. Along with these potential advantages, however, come a variety of new challenges as well. First, different species or cell types of interest must be able to coexist. Additionally, in many scenarios the relative abundance of each resident can impact the overall property of the community. Beyond coexistence and community composition, information processing and sharing is often essential to the types of complex, coordinated behavior that is required for many desired medical, industrial, and environmental applications.

My dissertation has centered around the design and implementation of two novel systems which address some of the challenges discussed above that must be overcome to realize the potential of synthetic microbial communities for use in technological applications. In the first system our goal was to develop a tool that can be used to enable coexistence and program community composition within a synthetic microbial community. We use

temperature as a modality to enable coexistence of two microorganisms, *Escherichia coli* and *Pseudomonas putida*, with different thermal niches and to further program the composition of this model synthetic bi-culture. Specifically, I developed two different approaches, referred to as a constant temperature regime and a cycling temperature regime. Employing a combination of wet-lab experiments and mathematical modeling, I showed that a variety of parameters such as temperature, cycle duration, etc. can be manipulated to achieve desired community compositions. Building on this work, I then used a mathematical framework developed by ecologists to explore design principles and specific mechanisms underlying the observed relationship between culture temperature and coexistence.

In the second system, I designed a novel synthetic microbial community with a distributed sensing and centralized reporting architecture that is enabled by what we have termed bacteriophage-mediated information transfer. Our goal is to explore a novel distributed sensing with centralized memory system architecture that is capable of addressing limitations of previously developed systems. A modular genetic circuit was developed that connects the input of an environmental signal of interest to activation of a lysogenic lambda bacteriophage which is used to transfer information about the sensing event from the sensor cell population to a reporter cell population. A variety of different ways to encode and store information were explored.

While seemingly different, the lines of work described above are connected by a common thread of developing generalizable and modular approaches for engineering synthetic microbial communities to deliver the potential advantages they offer in a variety of medical, industrial, and environmental applications. Synthetic microbial communities are capable of

performing complex and varied functions within these contexts and this dissertation is contributing to the rapidly growing body of research work for addressing the challenges that must be overcome to realize that potential.

Chapter 1: Background and Motivation

1.1 The genesis and general approaches of synthetic biology

The concept of engineering is nearly as old as human civilization. The modern definition of engineering revolves around the creative application of scientific principles and understanding of the natural world to design and generate structures, machines, and apparatus relevant to humankind (Britannica, 2020). As such, engineering has been, throughout its history and still, inextricably linked to the development and use of technology. In contrast, the discipline of molecular biology is relatively young. Generally accepted to have emerged in the 1930s, the distinct discipline of molecular biology arose from a fundamentally multidisciplinary attempt by biochemists, geneticists, microbiologists, virologists, and importantly physicists, to understand the fundamental physical and chemical processes of biological systems down to the molecular, and even atomic, level (Kellenberger, 2004). Between the 1930s and 1960s, in what is in retrospect a stunning demonstration of human ingenuity and scientific progress, the link between genes and proteins was established and the identity of DNA as the physical substance of genes was demonstrated, followed by the confirmation that DNA makes up the genetic material of bacteriophage, a virus which infects bacteria. Finally in 1953 the elucidation of the atomic structure of DNA by James Watson, Francis Crick, and Rosalind Franklin was published (Franklin & Gosling, 1953; Rheinberger, 2009; Watson & Crick, 1953). This absolute whirlwind of scientific progress led to two distinct concepts, one theoretical and the other technological,

whose consequences reverberate through human society to this day, and in all likelihood will continue to do so for the foreseeable future. First the theoretical, the formulation of the central dogma of molecular biology, which is now understood as the process by which the information in DNA is converted into RNA and then into protein, which are the workhorses of biological cells and systems. It is interesting to note that this fundamental process of biology, and life itself, is as much an informational process as a physical and energetic process. In any case, this understanding of the basic organizational hierarchy of biological systems, combined with the second innovation, the technological ability to isolate, manipulate, and characterize the physical substances of DNA, RNA, and proteins have given rise to the field commonly referred to as genetic engineering. Genetic engineering refers to the ability to modify DNA sequences and subsequently RNA and protein composition to produce desirable characteristics in biological systems. This combination of advances in both the theoretical understanding of biological systems combined with the technology to physically manipulate and characterize them at the atomic level have given rise to yet another new discipline, this time one that is only approximately 20 years old. This discipline, referred to as synthetic biology, endeavors to apply the philosophy and practices of engineering to biological systems as opposed to the more macromolecular structures that have been the purview of engineering throughout the course of its millennia old history. Synthetic biology attempts to leverage tried and true engineering concepts such as modularity, robustness, the engineering design cycle, etc. and apply them to biological systems to produce synthetic biological systems with properties that are desirable for human enterprise.

The first major success of synthetic biology dates to approximately 25 years before the concept of synthetic biology was formalized, with the cloning of the human insulin gene into the bacteria Escherichia coli to produce insulin for patients with diabetes (Goeddel et al., 1979). Descriptions of symptoms and treatments for diabetes mellitus can be dated back to ancient Egyptian, Chinese, and Indian civilizations (Quianzon & Cheikh, 2012). It was not until millennia later, in 1921, when animal insulin was isolated and used to treat human diabetes patients for the first time (Ladisch & Kohlmann, 1992; Quianzon & Cheikh, 2012). Over the next six decades the process for isolation and purification of insulin from animal sources was revised and refined but still struggled to meet the twin challenges of supply and demand and increasing immunological responses after repeated administration over time (Ladisch & Kohlmann, 1992). In 1978, in what is generally accepted as one of the first major successes for the biotechnology industry, the genes for the human insulin A and B chains were cloned into *E. coli* plasmids, expressed, purified, and combined to produce a molecule of (Sheth & Wang, 2018)insulin that is biochemically identical to human insulin despite having been manufactured by bacteria (Goeddel et al., 1979). This process, developed by researchers at Eli Lilly and Co. and Genentech, created a virtually unlimited supply of human insulin and exemplified one of the core biotechnology strategies that would be iterated upon, improved, and implemented consistently over the next two decades (Ladisch & Kohlmann, 1992).

This core biotechnology development strategy is conceptually straightforward, but not always trivial in implementation (Figure 1.1). First, researchers identify a gene of interest that codes for an RNA or protein product of perceived medical or technological value. This gene is then isolated, cloned, and inserted into a microbial expression vector (although mammalian cell culture has also become a common production vector). By orders of magnitude, the two most common microbial expression systems are Saccharomyces cerevisiae, the well-loved eukaryotic yeast microbe responsible for conversion of sugars into ethanol for wine and beer production, and Escherichia coli, the less well-loved (by the general public) prokaryotic commensal mammalian gut microbe. There are a multitude of reasons for the favoritism bestowed upon these two model systems, but the majority of those reasons boil down to a few broad characteristics. These microbes are relatively robust, grow extremely rapidly (with doubling times on the order of minutes to hours) at temperatures which are experimentally amenable, and by virtue of a research positive feedback cycle as much as anything else (wherein more research was done on *E. coli* because there were more data and tools available, which led to even more data and tools, etc.), have a

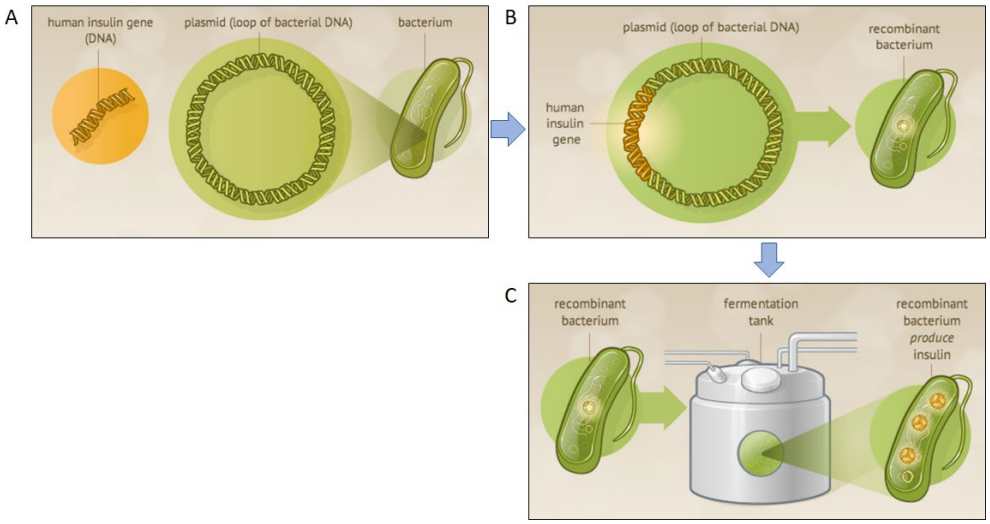


Figure 1.1 Gene of interest (insulin) is isolated. B. Gene of interest is introduced into a bacterial plasmid and the plasmid is transformed into a model organism for production. C. The production organism is cultured for production of the target molecule. (NIH)

wealth of genetic tools and physiological data available for researchers. Upon introduction of the gene of interest into the expression vector and model species of choice, the species can then be cultured at industrially relevant scales (often in culture volumes into the tens of thousands of liters) for production and subsequent extraction of the target molecule. As mentioned previously, this general approach has been applied again and again over the past four decades, albeit with wildly varying degrees of success.

1.2 Initial challenges in synthetic biology

Beginning with the initial success with insulin and a few other targets, interest in biotechnology has exploded in the last 40 years (Lazonick & Tulum, 2011). And despite what have certainly been a large number of success stories, expectations were soon tempered by reality. The nascent field encountered many challenges and complications resulting from a multitude of aspects of biology across several physical scales. First, while gene targets gradually became easier to identify with the development of bioinformatic tools such as metagenomics, isolation, and culture of many of the organisms that harbor the identified genes of interest can be excruciating. Recently, one microbe of interest was finally isolated after 2,013 days of continuous anaerobic culture and over 12 years of work (Aoki et al., 2014; Starr, 2019). Additionally, it is estimated that only approximately 1% of all microbial species sampled from natural environments have been culturable (Schloss & Handelsman, 2005). It is currently proposed that one of the major difficulties in culturing many microbes of interest in isolation stems from obligate symbiotic relationships between the bacteria of interest and other species they are found within their natural context, rendering many species incapable of being cultured in mono-culture (at least in the absence of specialized media) (Stewart, 2012). Even after

isolation of species enabling the harvest of sufficient quantities of high-quality DNA substrate of target genes for cloning purposes, the heterologous expression of foreign genes in model species offers many challenges of its own. Codon optimization, post-transcriptional and post-translational modification, and aggregation/inclusion bodies are just a few of the dimensions that must be optimized for expression of target genes in model organisms, often with very little in the way of a roadmaps for success (Burgess-Brown et al., 2008; Mukherji & van Oudenaarden, 2009; Purnick & Weiss, 2009; Slouka, Kopp, Spadiut, & Herwig, 2019). In the event these obstacles are overcome, it is also the case that many biosynthetic pathways of medical and industrial interest can be quite large (Weathers, Elkholy, & Wobbe, 2006). Attempting to express such a long biosynthetic pathway heterologously in target cells can introduce a metabolic burden such that the intended production cells display decreased robustness in addition to growth defects (Glick, 1995). Finally, at the population level, the industrially sized cultures can often be susceptible to invasion by foreign species of microbes as well as bacteriophage, resulting in culture collapse and significant cost increases.

1.3 A new approach for synthetic biology: synthetic microbial communities

1.3.1 Potential advantages associated with synthetic microbial communities

Significant progress has been made on developing techniques to address each of these fronts in the past decades. In the mid-to-late 2000s, however, a new strategy for the expression and optimization of biotechnological systems began to garner interest and attention. It was proposed that instead of using a single species culture approach to produce biomolecules of interest, it might be possible to engineer chemistry and coordination in a microbial community of multiple species to address many of the challenges that had presented

as obstacles for the synthetic biology field whose roots reach back to the late 1970s or further. For example, culturing fastidious or otherwise difficult to isolate species amongst their natural communities may provide the symbiotic relationships that are required for those organisms to thrive and grow (Stewart, 2012). Additionally, using a multi-species approach theoretically enables expression of target genes in their natural systems, which bypasses the plethora of challenges associated with heterologous expression in model systems (Glick, 1995). It can also potentially aid in diminishing the metabolic burden imposed upon production cells (Figure 1.2). In many cases, different steps of naturally occurring pathways are sourced from a variety of different organisms and all combined into a single biosynthetic pathway in the final model production organism. This is often at least partially responsible for imposing detrimental

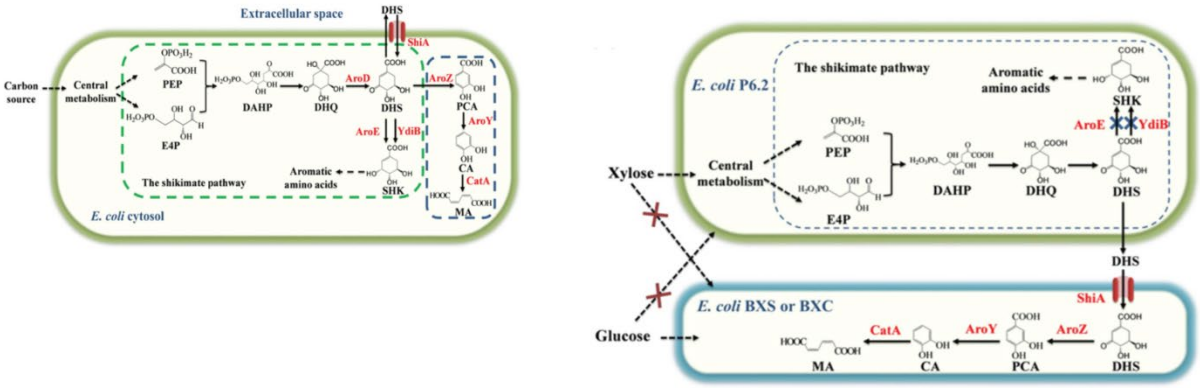


Figure 1.2 Adapted from Zhang et al. 2015. Division of a long metabolic pathway into two subpathways distributed amongst 2 different cell populations within a microbial community for production of a target molecule (H. Zhang, Pereira, Li, & Stephanopoulos, 2015).

metabolic burdens on production specialists (Glick, 1995). In a production community paradigm, the different pathway pieces can still be expressed individually in their natural hosts, with intermediates shuffled between specialists in the community. Additionally, the community approach enables the concept of specialization within the production scheme. Instead of forcing a single species to act as a generalist, performing all the biochemical

reactions and processes required for synthesis of a target, the community approach allows for the distribution and specialization of labor (Thommes, Wang, Zhao, Paschalidis, & Segrè, 2019). This enables optimization of specific reaction condition for specific pathway steps as well as reaction compartmentalization to minimize unwanted intermediates and carbon flux through non-target pathways (Brenner, You, & Arnold, 2008). Finally, there is a wealth of data, generated largely by ecologists, which suggests that biodiversity in naturally occurring microbial communities contributes to increased efficiency and biomass production, stability, and resistance to invasion by foreign species in comparison to monocultures (Cardinale et al., 2012; McGrady-Steed, Harris, & Morin, 1997; van Elsas et al., 2012). Each of these characteristics can be considered beneficial within a biotechnological context. In short, there are clearly many potential benefits to a microbial community-based production strategy, as long as the challenges they also present can be overcome.

1.3.2 Challenges specific to synthetic microbial communities

While microbial communities certainly offer the potential to overcome many of the challenges that have been encountered with single species approaches to biotechnology, they also present new challenges all their own. The first challenge can be summarized simply as coexistence. To access the potential benefits associated with synthetic microbial communities, the species intended as residents in the community must be able to coexist together over the relevant time frames (e.g. the length of culture time required for a production cycle). In addition to coexistence, the number of individuals of each different species in proportion to the total community population (also referred to as the relative abundance of each species) will often affect the metabolic properties of the community as a whole. For example, as long of the

metabolic properties of species A differ significantly from species B, a community that is 25% composed of species A and 75% composed of B would be expected to exhibit different metabolic properties as a whole than a community composed of 25% species B and 75% species A.

Within the field of synthetic biology there has been a severe shortage of tools and techniques developed to address challenges with enabling and regulating coexistence and relative abundance in synthetic microbial communities until recently (Scott et al., 2017; Stephens, Pozo, Tsao, Hauk, & Bentley, 2019). In addition, very little research has been done by synthetic biologists regarding the basic mechanisms which affect coexistence and relative abundance in microbial communities despite the fact that this kind of basic research is often required to identify mechanisms which can form the basis of tools and strategies to manipulate these coexistence and relative abundance in microbial communities.

Despite the lack of attention to these issues from synthetic biologists, these concepts have been the subject of a great deal of research by ecologists over the past century and a half, extending all the way back to Darwin's theory of natural selection (Darwin, 1859). One of the earliest paradigms for the study of coexistence by ecologists is commonly referred to as the competitive exclusion principle, commonly attributed to Gause but formalized by Grinnell in 1904 as "Two species of approximately the same food habits are not likely to remain long evenly balanced in numbers in the same region. One will crowd out the other" (Grinnell, 1904). This concept, which was simplified by Hardin into "complete competitors cannot coexist", was the subject of intense exploration and debate over the first half of the 20th century (Hardin, 1960). The competitive exclusion principle is predicted and supported by various mathematical

and theoretical models such as the Lotka-Volterra model of competition, but is often contradicted by empirical observations of naturally occurring communities, a juxtaposition famously articulated by Hutchinson in his 1960 publication 'The paradox of the plankton,' which contrasts the competitive exclusion principle with the many naturally occurring communities of aquatic plankton composed of multiple species of plankton which occupy very similar ecological niches and coexist over extended periods of time (Hutchinson, 1961). In this article, Hutchinson suggests that fluctuations in environmental conditions is one mechanism that could theoretically enable coexistence of competitors in natural communities (Hutchinson, 1961). Debate and investigation continued over the rest of the 20th century and in the year 2000 the mathematical ecologist Peter Chesson published a seminal paper which synthesized and clarified much of the coexistence theory from the previous century within a mathematically rigorous framework (Chesson, 2000).

Due in part to the complications and confounding factors arising in part from the immense diversity amongst the types of naturally occurring communities that are studied by ecologists, from terrestrial to aquatic environments at scales ranging from whales to microbes, investigation into coexistence maintenance mechanisms continues to this day, but much of the knowledge that has been generated by over a century of investigation is now sufficiently developed for application in synthetic biology as researchers attempt to engineer and manipulate many of the parameters that have been explored through ecological studies. Furthermore, it is the case that synthetic microbial communities in many cases provide a very useful model system to explore ecological questions due to their relatively fast generation times, small size, and easily scalable complexity, which allow highly complex communities to be

studied over many generations under experimentally amenable time frames and conditions (Scott et al., 2017).

In addition to the population level challenges with synthetic microbial communities discussed above, there are molecular level challenges that arise in attempting to engineer

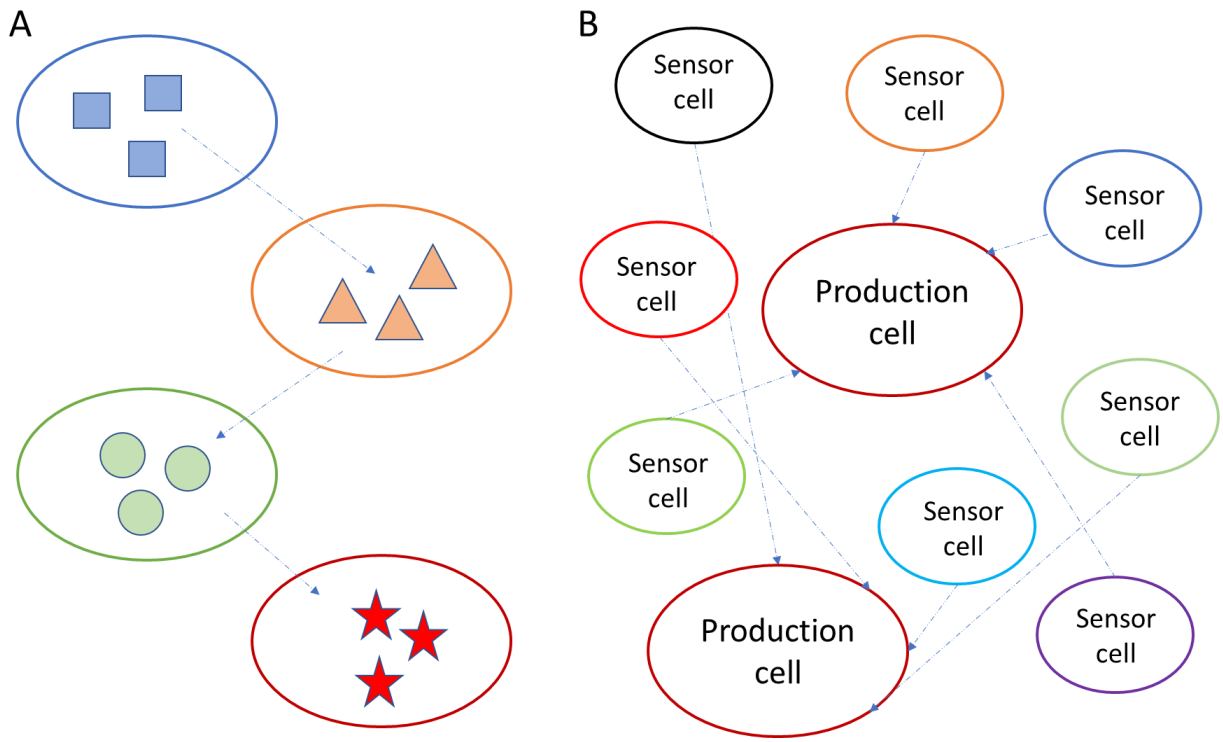


Figure 1.3 A. Four different cell types in a microbial production community encode different parts of a biosynthetic pathway (four large differently colored ovals). Pathway intermediates (represented by blue squares, orange triangles, and green circles) are passed between cells in the production pathway to finally produce the target product, represented by red stars. B. Production of a target compound in the production cells (red ovals) is controlled by signals sent to the production cells from sensor cells in the community.

synthetic microbial communities for biotechnology applications as well. At the most basic level, community resident species must be capable of sharing the compounds upon which they are meant to act, and at a more complex (and arguably more important) level, community species must somehow be able to coordinate their activity along relevant temporal and spatial scales to enable the complex community behaviors that are desired for biotechnology applications (**Error! Reference source not found.**) (DeLisa, Valdes, & Bentley, 2001; Dunn & Handelsman,

2002; Humphries et al., 2017; Scott et al., 2017; Silva & Boedicker, 2019; von Bodman, Willey, & Diggle, 2008). In order to enable this coordination and cooperation within the community, the most important commodity that must be shared is information.

Species must be able to regulate the correct chemical and biophysical processes at the appropriate times in response to the desired environmental and internal conditions of not just one but often multiple different types of individual species and cell types. It is this level of coordination and cooperation at which one of the cutting edges of synthetic biology in microbial communities currently lies.

1.3.3 Overcoming the challenges: Inspiration from naturally occurring microbial communities

While these challenges appear perhaps at the outset to be daunting, there are thankfully a wealth of naturally occurring microbial communities which exhibit exquisite coordination across time and space to which we can turn for inspiration. For example, the light generating organ of the Hawaiian squid Euprymna scolopes forms an association with the Gram-negative luminous bacterium Vibrio fischeri that lasts over many generations of V. fischeri and for the entirety of

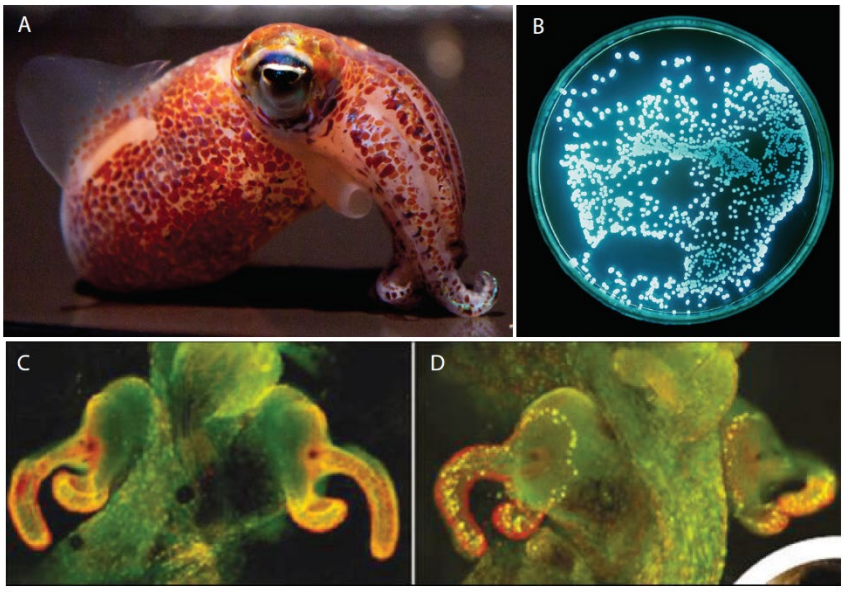


Figure 1.4 A. The squid Euprymna scolopes *B. The symbiotic bacterium* Vibrio fischeri *with which the squid forms a lifelong association C-D The light producing organ of* Euprymna scolopes *(Olsen, Choffnes, & Mack, 2012).*

the lifetime of the squid (Figure 1.4) (Olsen et al., 2012). The bacteria of the light organ emit luminescence that mimics the aesthetic of moonlight and starlight filtering through ocean waters, thus camouflaging the nocturnal squid from predators below (Nyholm & McFall-Ngai, 2004; Olsen et al., 2012). The association between the light organ of the squid of V. fischeri begins within hours of hatching and occurs only with the V. fischeri bacterium. The association proceeds in stages, and at each stage the specificity between host and symbiont is increased. Upon completion of this colonization process, the bacterium drives development of the tissues within the light organ, inducing a morphology change in the squid tissue that enables a transition from a physiology that supports colonization to one that promotes long term maintenance of a long term exclusive association (Nyholm & McFall-Ngai, 2004; Olsen et al., 2012). While the exact molecular mechanisms of these interactions are still under investigation, it is certain that the exchange of information in some form, possibly some type of biomolecules or morphogens, must occur for this type of coordination to occur.

Another example of complex coordination and communication in microbial communities comes from the microbial communities associated with the roots of plants, commonly referred to as the plant rhizosphere (Figure 1.5). The rhizosphere of a plant is generally a complex microbial community composed of species that exchange energy and nutrients with the plant in a symbiotic relationship (Ferguson et al., 2010; Long, 2001; Olsen et

al., 2012). It is known that partnerships between plant roots and the diverse species of the rhizosphere community are established in part through chemical and genetic cross talk in which chemical signals can be released by plants to attract specific microbes, which upon association release additional signals that alter gene expression in the plant in ways that can affect plant morphology and biochemistry (Ferguson et al., 2010; Long, 2001; Olsen et al., 2012). So while

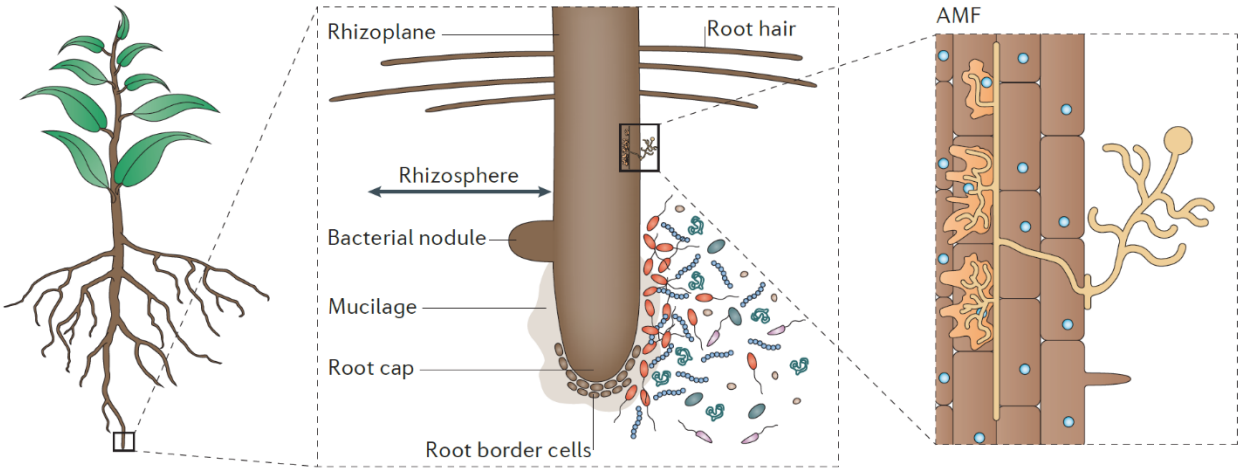


Figure 1.5 A cartoon of the associated formed between plant roots and the microbiota of the rhizosphere. Left, a macroscopic view of the full plant. Middle, a close up of the associations between the microbiota (in red, blue, green, and yellow and the plant tissue (brown). Right, infiltration of the plant tissue by arbuscular mycorrhizal fungi (AMF) (Philippot, Raaijmakers, Lemanceau, & van der Putten, 2013).

the challenges of coordinating complex behaviors between different species in a synthetic microbial community are certainly not trivial, it is possible to look to naturally existing microbial communities to see that it is indeed possible, and also to consider that strategies used in these natural communities and whether such strategies might be appropriate for synthetic communities or if they inspire new strategies that would be appropriate.

1.4 Dissertation overview

The research and ideas discussed above provide an appropriate introduction to the underlying motivations and goals of my thesis work over the past six years. The fundamental goal of synthetic biology has been to apply engineering philosophies and strategies to the

practice of molecular biology to develop biotechnology applications of biological systems to address environmentally, industrially, and medically relevant current challenges. A recently evolved paradigm of synthetic biology expands the approach of using single species microbial cultures for implementation of desired biological systems to using synthetic microbial communities which offers a wide variety of potential improvements over traditional single species approaches alongside a host of new challenges. The projects I have worked on in my thesis work have been specifically designed to leverage ecological research in addition to engineering and molecular biology techniques to develop strategies to enable coexistence and program relative species abundance within synthetic microbial communities and to introduce a new information sharing paradigm within synthetic microbial communities to enable communication and coordination in microbial communities using a variety of both wet lab techniques as well as computational and mathematical models. In Chapter 2 I present data from a novel regulatory system we developed to enable species coexistence and program relative species abundance in a synthetic microbial community. Rationally designed temperature regimes are used to exert control over relative species abundance by differentially manipulating the co-culture growth rates of community members. In Chapter 3, I more deeply explore the relationship between temperature, growth rates, and population dynamics using a mathematical framework developed by ecologists to probe mechanisms of coexistence maintenance in natural communities. Finally, in Chapter 4 I present a novel synthetic microbial community designed to be capable of sensing and remembering a variety of environmental stimuli to which it is exposed. We accomplish this by implementing a novel distributed sensing system architecture paired with a centralized memory component enabled by bacteriophage-

mediated information transfer. Together these projects represent a significant contribution towards developing the ability to better engineer and program synthetic microbial communities for biotechnological applications.

Chapter 2: Temperature Regulation as a Tool to Program Synthetic Microbial Community Composition

The following chapter is largely based on our manuscript currently in revision for publication in Biotechnology and Bioengineering.

2.1 Introduction

Microbial systems have been used for decades to perform functions with industrial, environmental, and medical relevance (Bouwer & Zehnder, 1993; Goeddel et al., 1979; Houde, Kademi, & Leblanc, 2004). Traditionally, a single microbial species is cultured for a specific functionality. If necessary, genetic modifications are made to this organism to enable novel capabilities. As the field surrounding this paradigm has progressed, however, challenges faced by this monoculture approach have become increasingly apparent. Metabolic versatility and complexity in a single species is limited by factors such as metabolic burden, a lack of or limited intracellular compartmentalization, and energetic tradeoffs (Litchman, Edwards, & Klausmeier, 2015; K. Zhou, Qiao, Edgar, & Stephanopoulos, 2015). At the population level, monocultures are susceptible to invasion and subsequent population collapse by agents such as foreign bacteria, fungi, and bacteriophages. One approach to address these challenges is utilization of synthetic microbial communities as opposed to monocultures for biotechnology applications. Microbial communities enable division of labor and specialization within a population,

decreasing metabolic burden and enabling reaction compartmentalization. More diverse microbial communities can also exhibit increased stability (Yurtsev, Conwill, & Gore, 2016) and resistance to invasion by foreign agents (van Elsas et al., 2012), as well as increased biomass production (McGrady-Steed et al., 1997) and nutrient utilization (Ptacnik et al., 2008) in certain contexts.

While the potential advantages offered by a microbial community over a monoculture are clear, community-based approaches face challenges of their own. One of the most basic challenges is control of community composition. Community composition can have different meanings in different ecological contexts, but here is used to refer specifically to the relative abundance of individual species within the community. In a microbial community in which each species exhibits a specific and unique function within the community, the metabolic functionality of the community as a whole can depend greatly on the relative abundance of each species within the community (B. B. Hsu et al., 2019; S. A. Scholz, Graves, Minty, & Lin, 2018; H. Zhang et al., 2015). Despite significant advances in many related research topics, progress in control of synthetic microbial community compositions has been slow, likely due at least in part to a dearth of tools capable of dynamically modulating relative species abundance. The most commonly used technique is manipulation of inoculation ratio (Minty et al., 2013; Shong, Jimenez Diaz, & Collins, 2012; Strickland, Lauber, Fierer, & Bradford, 2009), but this approach is not capable of dynamic control and due to differences between intrinsic growth rates and interspecific interactions between community residents the community composition often quickly shifts away from this initial condition.

While synthetic biology researchers are starting to develop new approaches to address the aforementioned challenges, mechanisms that enable species coexistence and modulate

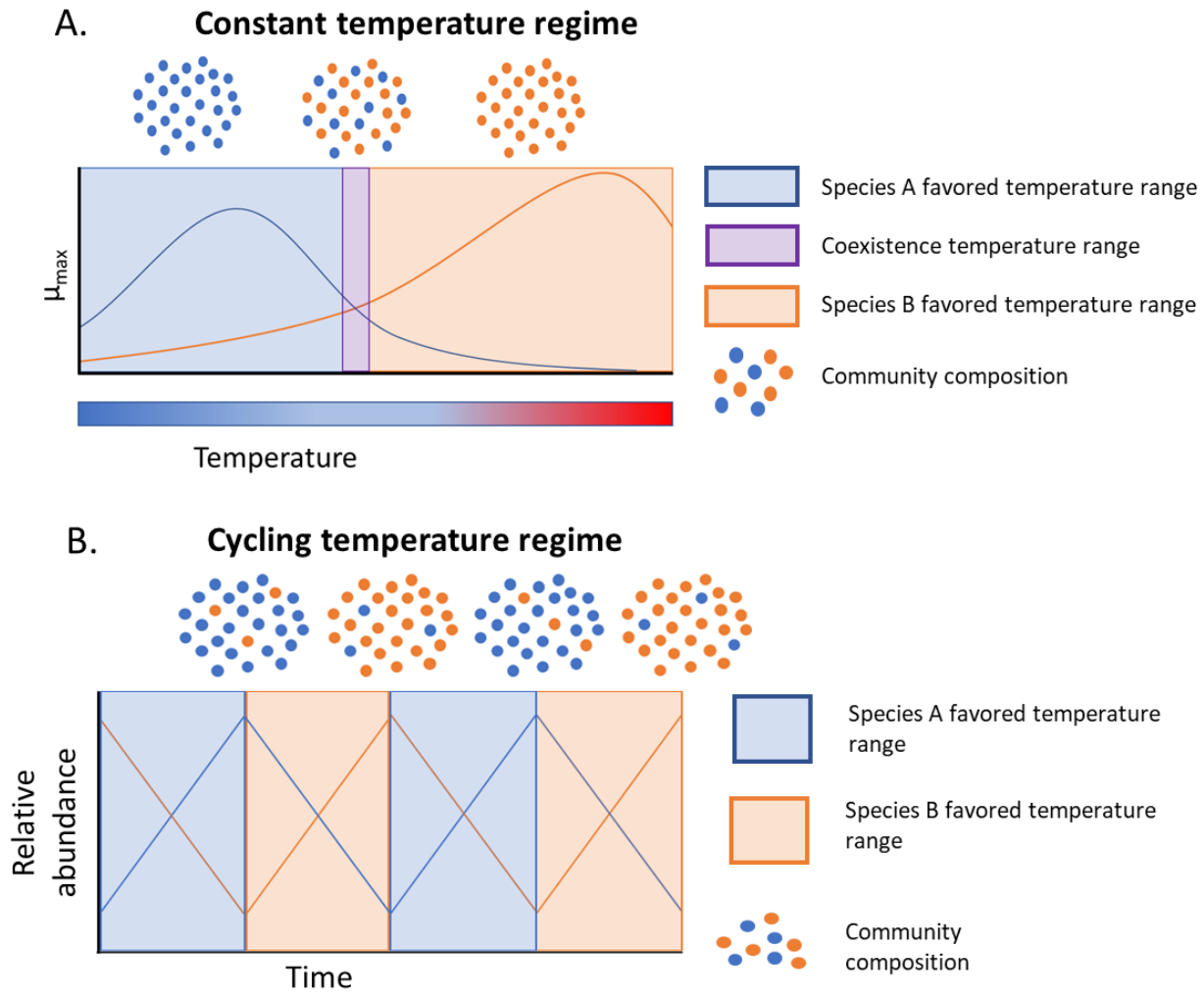


Figure 2.1 Two approaches for rationally regulating synthetic microbial community compositions. A) Under what we refer to as a constant temperature regime, a synthetic bi-culture is kept at a constant temperature throughout the duration of the culture lifetime. The selected temperature determines whether a monoculture persists or co-existence of the two microbes can occur. B) Under a cycling temperature regime a bi-culture is grown in repeated cycles of high and low temperatures, enabling co-existence of the two microbes over an extended period of time. Both the temperatures and the time intervals spent at each temperature can be manipulated to achieve desired outcomes.

relative abundance over time in complex naturally occurring communities have been a major theme in ecological study for almost 100 years. The current state of the field proposes that all mechanisms that maintain and modulate species diversity operate through one of two broad classes, equalizing mechanisms and stabilizing mechanisms (Chesson, 2000). Equalizing

mechanisms lessen the relative fitness differences between species (which can be more simply conceptualized as growth rates, although in truth this is a slight oversimplification), and stabilizing mechanisms increase the niche differences between species, which is to say they decrease the extent to which species compete directly with each other(Chesson, 2000). According to Chesson's theory, equalizing mechanisms can support coexistence by decreasing the difference between growth rates, but stabilizing mechanisms are required for coexistence. This makes intuitive sense in that, if there is a difference between the growth rates of two species, no matter how small, eventually that difference will lead to competitive exclusion of the slower growing species in the absence of additional compensatory mechanisms. Therefore, any environmental variables which differentially affect growth rates or niche partitioning would be expected to influence the ability of species to coexistence and/or community composition.

In this work we have applied concepts from the ecological study of the mechanisms enabling maintenance of species coexistence to a synthetic microbial community in order to develop tools that are capable of dynamically modulating relative species abundance. The modality we have implemented is temperature regulation. Using a synthetic community consisting of two model microorganisms Escherichia coli and Pseudomonas putida, which occupy different but partially overlapping thermal niches, we demonstrate that temperature regulation is a tool that can be used to enable coexistence and program the community composition. Specifically, we develop two temperature-mediated regimes. First, using a constant temperature for the co-culture, we are able to identify a temperature window that allows co-existence of the two microbes, albeit exhibiting only a small range of achievable compositions. Next, we develop a novel regime of cycling between a high temperature favoring

E. coli and a low one favoring *P. putida*, which enables co-existence of the two microorganisms in a highly tunable manner.

2.2 Materials and methods

2.2.1 Microbial strains and cultivation

E. coli K12 substr. MG1655 constitutively expressing a chromosomally integrated YFP construct was generated by P1 transduction as described by Thomason et al. (Thomason, Costantino, & Court, 2007). Strain Y2 from Kerner et al. (Kerner, Park, Williams, & Lin, 2012) was used as the source of the YFP cassette. The *P. putida KT2440* strain constitutively expressing mCherry was a generous gift from Dr. Esteban Martinez-Garcia of the Victor de Lorenzo lab. Addition of an mCherry cassette to *P. putida KT2440* was performed via Tn7 transposon assisted cloning (Koch, Jensen, & Nybroe, 2001; Martínez-García, Calles, Arévalo-Rodríguez, & de Lorenzo, 2011). Unless otherwise noted, all cultures were grown in minimal M9 media (200 mL 5x M9 salts [34 g/L Na2HPO4 anhydrous, 15 g/L KH2PO4, 2.5 g/L NaCl, 5 g/L NH4Cl], 20 mL 20% glucose, 2 mL 1M MgSO4, 100 uL 1M CaCl2, 780 mL sterile deionized H2O; autoclave all components separately then mix under sterile conditions).

2.2.2 Visualization of strains via confocal microscopy

50 uL of log phase bi-culture was pipetted onto a slide and a coverslip was placed on top. Confocal fluorescence microscopy was performed using an upright Olympus FV1200 confocal microscope equipped with a 60x objective and 405, 488, 515, 543, and 635 lasers. Images were exported and analyzed with ImageJ.

2.2.3 Quantification of growth rate in co-cultures

E. coli K12 substr. MG1655-YFP and *P. putida KT2440*-mCh were seeded from -80 ˚C cryostocks and grown in monoculture overnight to stationary phase in 14 mL Corning Falcon® test tubes (polypropylene test tube, round bottom, 17x100mm, 14mL, graduated, with clear snap cap, Sterile, 25 per Pack) in a 2 mL volume of M9 minimal media. Overnight cultures were diluted 1:100 into fresh M9 media and grown into exponential growth phase (OD600 ~0.4-0.6). The cell density of the two cultures of exponential phase cells were normalized to each other using OD600 and then diluted 1:100 into fresh M9 media in a Greiner Bio-one CELLSTAR™ 96 well μClear flat bottomed microplate with a 200 uL final volume. Strains were grown in monoculture and co-culture in triplicate (2 uL of the desired strain was added for monocultures, and 1 uL each of both strains was added for co-cultures to 198 uL M9 media). These nine wells were surrounded by wells filled with 200 uL sterile deionized H2O to inhibit evaporation from experimental wells. The lid was treated with a mixture of 20% ethanol + 0.5% Triton X-100 to avoid condensation formation on the lid (pour enough mixture to completely cover the bottom of the lid, let sit 5 minutes, pour off and let air dry). The lid was taped on using Fisherbrand™ labeling tape. The plate was incubated in a Biotek Synergy H1 platereader for 24 hours with plate reads every 10 minutes and continuous orbital shaking at 282 cpm. At each timepoint, reads at 600 nm wavelength, Excitation: 510 nm Emission: 540 nm, and Excitation: 585 nm Emission: 620 nm were taken of each well (these wavelengths were empirically determined to maximize the specific signal and minimize crosstalk between the two channels for our constructs, media, and platereader; data not shown). The co-culture growth rates from the co-

culture growth curves were calculated by fitting the early exponential growth phase to an exponential function using Microsoft Excel.

2.2.4 Constant temperature regime: culture growth and community composition quantification

E. coli K12 substr. MG1655-YFP and *P. putida KT2440*-mCh were seeded from -80 ˚C cryostocks and grown in monoculture overnight to stationary phase in 14 mL Corning Falcon® test tubes (polypropylene test tube, round bottom, 17x100mm, 14mL, graduated, with clear snap cap, Sterile, 25 per Pack) in a 2 mL volume of M9 minimal media. Overnight cultures were diluted 1:100 into fresh M9 media and grown into exponential growth phase (OD600 ~0.4-0.6). The cell density of the two cultures of exponential phase cells were normalized to each other using OD600 and then diluted into 14 mL Corning Falcon® test tubes with 2 mL fresh M9 at three different inoculation ratios; 1:10, 1:1, and 10:1 *E. coli* : *P. putida* (2 uL *E. coli* + 18 uL *P. putida*, 10 uL *E. coli* + 10 uL *P. putida*, and 18 uL *E. coli* + 2 uL *P. putida*). Each inoculation ratio was performed in triplicate producing a total of nine cultures. Cultures were incubated in a Lab-Line Instruments Model No. 3528 incubator with a New Brunswick Scientific C1 platform shaker placed inside for shaking set at a speed of 50 (no units are indicated on the shaker). The temperature was monitored throughout the lifetime of each experiment using an Omega OM-91 portable temperature data logger (see Supplemental Information). Each culture was passaged 1:100 into 2 mL fresh M9 media in 14 mL Corning Falcon® test tubes twice daily (8 hour and 16-hour growth periods). At each passaging time point, each culture was vortexed for 5 seconds and a 1 uL sample was taken for quantification of community composition via flow cytometry. Passaging was performed after vortexing.

2.2.5 Cycling temperature regime: culture growth and community composition quantification

E. coli K12 substr. MG1655-YFP and *P. putida KT2440*-mCh were seeded from -80 ˚C cryostocks and grown in monoculture overnight to stationary phase in 14 mL Corning Falcon® test tubes (polypropylene test tube, round bottom, 17x100mm, 14mL, graduated, with clear snap cap, Sterile, 25 per Pack) in a 2 mL volume of M9 minimal media. Overnight cultures were diluted 1:100 into fresh M9 media and grown into exponential growth phase (OD600 ~0.4-0.6). The cell density of the two cultures of exponential phase cells were normalized to each other using OD600 and then diluted into a 14 mL Corning Falcon® test tube with 2 mL fresh M9 at a 1:1 inoculation ratio (10 uL *E. coli* + 10 uL *P. putida*). For the long-time interval cycling temperature program experiments, the culture was incubated at 37 ˚C for 16 hours in a New Brunswick Scientific Excella E24 Incubator Shaker with 225 rpm shaking, and at 31 ˚C for 8 hours in a Lab-Line Instruments Model No. 3528 with 225 rpm shaking. For the short-time interval cycling temperature experiments, a culture was prepared as above and incubated for approximately 72 hours to allow adaptation to culture conditions. The culture was then incubated at 27 ˚C until the community composition reached a 1:1 ratio, and then was incubated at the indicated temperatures for the indicated times. After each growth period, cultures were vortexed for 5 seconds and a 1 uL sample was taken for quantification of community composition by flow cytometry, and then the culture was passaged 1:100 into 2 mL of fresh M9 media in a new 14 mL Corning Falcon® test tube.

2.2.6 Quantification of community composition via flow cytometry

Samples taken from co-cultures were diluted to ~106 cell/mL in 1x PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na2HPO4, 1.8 mM KH2PO4; pH 7.4) and run on an Applied Biosystems Attune acoustic focusing cytometer. 300 uL of sample was acquired by the device for each run and 10,000 events were recorded. The instrument configuration was as follows: (Threshold (x1000): FSC: 10; SSC: 10; BL1: 10; BL2: 10; BL3: 10; BL4: 10; RL1: 10; RL2:10, Voltage (mV): FSC: 3,000; SSC: 3,500; BL1: 2,400; BL2: 1,800; BL3: 2,550; BL4: 2,550; RL1: 2,900; RL2: 1,950). Because this machine lacks a laser with proper wavelengths for detection of our mCherry construct, yellow fluorescence positive cells were classified as *E. coli* K12 substr. MG1655-YFP and yellow fluorescence negative cells were classified as *P. putida KT2440*-mCh.

2.2.7 Modeling of population dynamics under the cycling temperature regime

For the cycling temperature regime, consider each cycle consists of a time interval t^H at a high temperature favoring *E. coli* and then another time interval t^L at a low temperature favoring *P. putida*. Two types of models were developed. In the first one, simple exponential growth is assumed.

$$\frac{dE}{dt} = r_E E$$

$$\frac{dP}{dt} = r_P P$$

where variables E and P are *E. coli* and *P. putida* cell densities respectively; r_E and r_P, their specific growth rates respectively, are model parameters. We further denote r_E^H and r_P^H as the specific grow rates of *E. coli* and *P. putida* respectively at the high temperature; r_E^L and r_P^L the

specific grow rates of *E. coli* and *P. putida* respectively at the low temperature. Note that $r_P^H < r_E^H, r_P^L > r_E^L$, and we define $\Delta r^H \equiv r_E^H - r_P^H, \Delta r^L \equiv r_P^L - r_E^L$.

The above ODE's have simple analytical solutions and the two cell densities can be expressed as explicit functions of time:

$$E(t) = \begin{cases} E_0 \cdot e^{r_E^H \cdot t}, & 0 \le t \le t^H \\ E_0 e^{r_E^H \cdot t^H} \cdot e^{r_E^L (t - t^H)}, & t^H \le t \le t^H + t^L \end{cases}$$

$$P(t) = \begin{cases} P_0 \cdot e^{r_P^H \cdot t}, & 0 \le t \le t^H \\ P_0 e^{r_P^H \cdot t^H} \cdot e^{r_P^L (t - t^H)}, & t^H \le t \le t^H + t^L \end{cases}$$

where E_0 and P_0 are the *E. coli* and *P. putida* cell densities respectively at the beginning of the cycle.

To maintain the bi-culture stably, the composition at the end of each cycle needs to return to its value at the beginning of the cycle. Therefore, we require the following:

$$\frac{E(t^H + t^L)}{P(t^H + t^L)} = \frac{E_0 e^{r_E^H \cdot t^H} e^{r_E^L \cdot t^L}}{P_0 e^{r_P^H \cdot t^H} e^{r_P^L \cdot t^L}} = \frac{E_0}{P_0} \cdot e^{(r_E^H - r_P^H) t^H - (r_P^L - r_E^L) t^L} = \frac{E_0}{P_0}$$

It follows that:

$$(r_E^H - r_P^H) t^H - (r_P^L - r_E^L) t^L = 0$$

Therefore, we have the relationship:

$$\frac{t^H}{t^L} = \frac{r_P^L - r_E^L}{r_E^H - r_P^H} = \frac{\Delta r^L}{\Delta r^H}$$

The ratio of the two species averaged over the whole cycle can be calculated as follows:

$$\left(\frac{E}{P}\right)_{average} = \frac{\int_0^{t^H}\left(\frac{E_0 \cdot e^{r_E^H \cdot t}}{P_0 \cdot e^{r_P^H \cdot t}}\right)dt + \int_{t^H}^{t^H+t^L}\left[\frac{E_0 e^{r_E^H \cdot t^H} \cdot e^{r_E^L (t-t^H)}}{P_0 e^{r_P^H \cdot t^H} \cdot e^{r_P^L (t-t^H)}}\right]dt}{t^H + t^L}$$

Simplifications of this expression, including utilization of the relationship above between time intervals and growth rate differences required for maintaining the bi-culture, ultimately lead to the following (Eq. 3.4.2):

$$\left(\frac{E}{P}\right)_{\text{average}} = \frac{E_0}{P_0} \cdot \left[\frac{\Delta r^L\left(e^{\Delta r^H \cdot t^H} - 1\right) + \Delta r^H\left(e^{\Delta r^L \cdot t^L} - 1\right)}{\Delta r^H \cdot \Delta r^L \cdot (t^H + t^L)}\right]$$

We also developed a second type of mathematical model where logistic growth is assumed.

$$\frac{dE}{dt} = r_E E(1 - \frac{E}{K_E})$$

$$\frac{dP}{dt} = r_P P(1 - \frac{P}{K_P})$$

where r_E and r_P are the maximum specific growth rates of *E. coli* and *P. putida* respectively; K_E and K_P are the corresponding carrying capacities. A Mathlab script (see Supplemental information) was developed to simulate repeated cycles of cell growth at a high temperature and then low temperature using these two un-coupled ordinary differential equations.

2.3 Results

2.3.1 Overall approach

Our objective is to develop new tools for regulating synthetic microbial community compositions, by leveraging temperature-dependent growths of distinct desired community

members. As illustrated in Figure 2.1, for a simple community consisting of two members exhibiting different growth phenotypes at various temperatures, we propose two temperature-mediated regimes for achieving co-existence and regulating community compositions. In the first regime (Figure 2.1 A), a constant temperature is used for growing the bi-culture. It is expected that if a low temperature is selected, the microbe favored at the temperature (i.e. exhibiting a higher growth rate) would dominate, and vice versa at a high temperature. We hypothesize, however, for a small range of intermediate temperatures, the two microbes could co-exist. In the second regime (Figure 2.1 B), the bi-culture is cycled between a low temperature and a high one, which favor each of the two microbes respectively, and co-existence of the two microbes are maintained over an extended period of time. Operating parameters include the high and low temperatures, and the time durations the bi-culture spends at each temperature. By manipulating these parameters, one would be able to regulate the average community composition and its variation.

2.3.2 Design and implementation of a model microbial bi-culture

To explore the temperature regulation regimes proposed above, we first established a model synthetic microbial system consisting of two well-characterized bacterial species. One is the soil bacterium Pseudomonas putida KT2440 (Belda et al., 2016). *P. putida* is found throughout the natural environment, predominantly in soil and plant rhizospheres (Belda et al., 2016; Thomas, Okamoto, Bankowski, & Seto, 2013). This bacterium is a mesophilic organism, growing optimally at relatively moderate temperatures (Munna, Zeba, & Noor, 2016), and has an extremely versatile metabolism which has earned it a reputation as an excellent bioremediation agent (Belda et al., 2016). In this study we use a strain of *P. putida* which has

been tagged with a constitutively expressed chromosomal copy of mCherry fluorescent protein (mCh) (Martínez-García et al., 2011; Rochat, Péchy-Tarr, Baehler, Maurhofer, & Keel, 2010). The second bacterial species in our system is *E. coli K12 MG1655* (Blattner et al., 1997). One of the most commonly used model organisms in microbiology, *E. coli* can be either commensal or pathogenic gut colonizers of humans and other mammals (depending on the strain), growing optimally at relatively higher temperatures compared to *P. putida*. In this study we used a strain of *E. coli K12 MG1655* which has been tagged with a constitutively expressed chromosomal copy of yellow fluorescent protein (YFP). Microbial species have naturally evolved to modulate their growth rate in response to temperature, often with optimal growth occurring within a small range of temperature. We consider this pair of bacteria as a model system of microbial communities composed of members occupying different but partially overlapping thermal niches. They can be grown together in a bi-culture in minimal M9 medium.

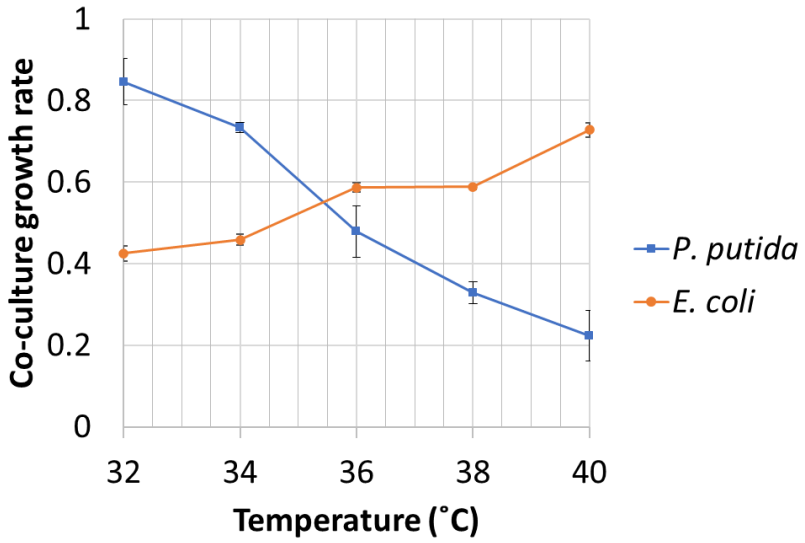


Figure 2.2 Co-culture growth rates of each bacterium in co-cultures with different temperature profiles. E. coli K12 substr. MG1655 grows optimally at relatively higher temperature and Pseudomonas putida KT2440 grows optimally at relatively lower temperatures. Error bars are standard deviation from three replicates.

We verified expression of the fluorescent proteins in each strain with confocal microscopy (Appendix Figure C.1 A) and flow cytometry (Appendix Figure C.1 B). We then quantified the growth phenotype of each species in co-culture at various temperatures.

Particularly, we determined the growth rates at different temperatures (Figure 2.2 Co-culture growth rates of each bacterium in co-cultures with different temperature profiles. E. coli K12 substr. MG1655 grows optimally at relatively higher temperature and Pseudomonas putida KT2440 grows optimally at relatively lower temperatures. Error bars are standard deviation from three replicates.Figure 2.2, referred to as temperature profiles hereafter), using growth curves derived from population level fluorescence from each species in bi-cultures. The temperature profiles indicate that for each species there exists a temperature range in which the species would be expected to have a growth advantage over the other (when the temperature is less than ~35.5 °C, *P. putida* μmax > *E. coli* μmax and when the temperature is greater than ~35.5 °C then *P. putida* μmax < *E. coli* μmax).

2.3.3 Effects on community composition of a constant temperature program

We first explore the constant temperature regime by investigating the community population dynamics over time at various selected temperatures. In this series of experiments, the temperature was maintained at a constant value (±0.2 ˚C, Appendix Figure D.1) throughout the entire course of an experiment, and co-cultures were diluted 1:100 into fresh media twice daily (9 h and 15 h growth periods). Each experiment was conducted for times ranging from 60 to 400 hours depending upon the observed dynamics. Community composition was quantified via flow cytometry with a measurement taken at each dilution time point. The dynamics of the bi-culture is shown in Figure 2.3 A for three representative temperatures, 35.5 ˚C, 36.5 ˚C, and 39.5 ˚C; the full set of data for 10 temperatures in the range of 32-39.5 ˚C are provided in detail in Appendix Figure B.1 and summarized in Figure 2.3 B. It was found that at 35.5 °C and lower, the community was dominated by *P. putida* when the culture composition reached equilibrium

(which is defined herein as changing no more than ±15% over at least 3 consecutive time points), regardless of imposed initial ratios of the two bacteria (for example, see Figure 2.3 A top row). On the other hand, at temperatures above ~39 ˚C the community was dominated by *E. coli* (for example, see Figure 2.3 A bottom row). Interestingly, at temperatures between 35.5 ˚C and ~39 ˚C, coexistence of the two bacteria was observed (for example, see Figure 2.3 A middle row). It is worth noting that the community composition at equilibrium of this bi-culture

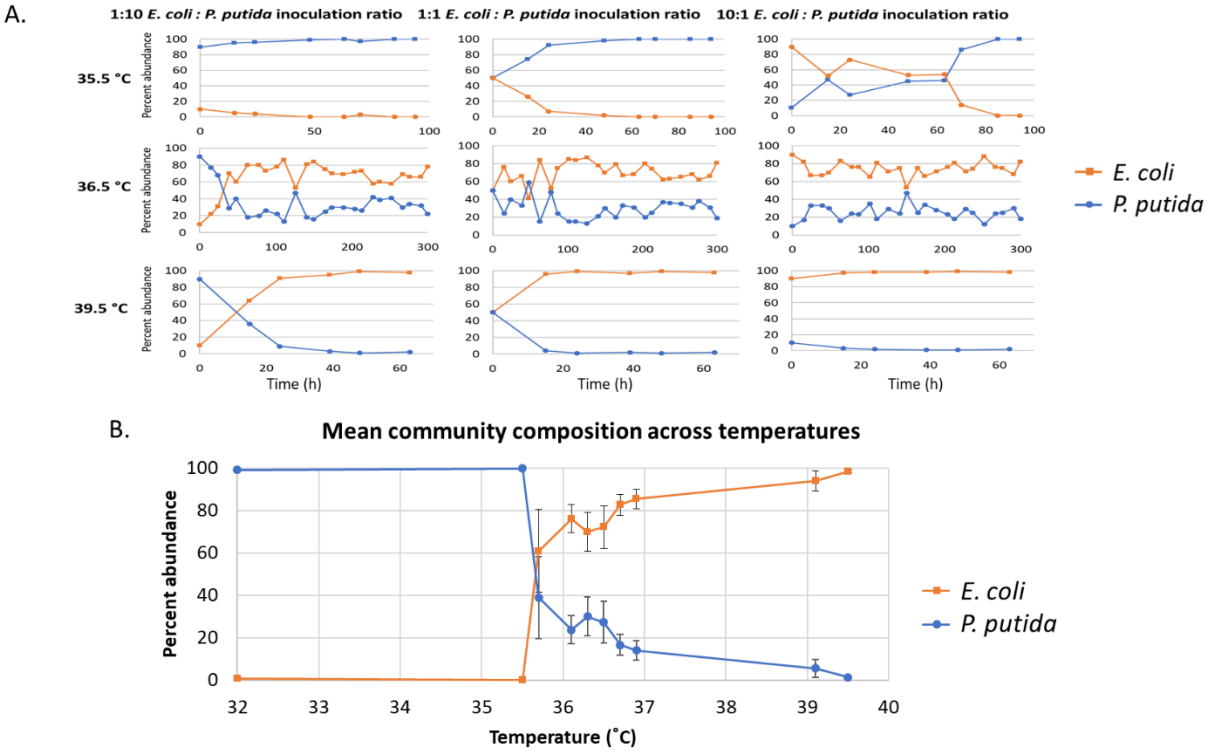


Figure 2.3 Constant temperature regimes can be defined that result in competitive exclusion of either species as well as coexistence. A) Representative graphs of species relative abundance over time of bi-culture with varying inoculation ratios grown at different constant temperatures. Each condition has three replicates B) Mean community compositions at each temperature. Error bars are standard deviation of the mean community compositions across all time points from all replicates at all inoculation ratios.

depends on the temperature in a highly nonlinear manner. Specifically, the equilibrium composition changed very gradually in the relatively wide temperature window of 35.7-39.1 °C; whereas this property changed drastically over a much narrower temperature window of 35.5-35.7 °C. These differences have direct implications for practical applications where a constant

temperature is employed. Given the resolution of temperature control (e.g. ±0.2 ˚C in our laboratory set-up), it is practically impossible to achieve a specific desired community composition in the small temperature window where the composition is highly sensitive to fluctuations in temperature. In contrast, in the temperature window where composition changes are milder, it is indeed possible to select a temperature to maintain the desired community composition. However, it is important to note that the range of community compositions achievable in this temperature window (~70-95% *E. coli* at 36.1-39.1 °C) is rather limited.

2.3.4 Regulating community composition using a cycling temperature program

We next explored the effect of temperature on community composition using what we refer to as a cycling temperature regime. In this approach, two distinct temperatures are selected so that the higher temperature favors the growth of one species and the lower one the other species. With these two temperatures selected, the bi-culture is incubated in repeated cycles, of which each starts with growth at the higher temperature for a specific time internal and then switches to growth at the lower temperature for another specific time interval. The key parameters we can manipulate are the two time intervals spent at the two temperatures, respectively. When they are selected properly, this scheme enables the two species to coexist. Moreover, the composition of the bi-culture can be tuned as desired and it is determined by a combination of the initial composition and the two time intervals. This can be illustrated analytically if we use a simple mathematical model where each species grows exponentially, which is a reasonable assumption when bacterial cells grow at low densities (i.e. far away from the carrying capacities). Specifically, the two species can be stably maintained provided that the

two time intervals are chosen such that the following relationship is satisfied (see Materials and Methods for details):

$$\frac{t^H}{t^L} = \frac{\Delta r^L}{\Delta r^H} = \frac{r_P^L - r_E^L}{r_E^H - r_P^H}$$

where t^H and t^L are the time intervals the culture spends at the high temperature and low temperature respectively in each cycle. r_P^L and r_E^L denote the specific growth rates of *P. putida* and *E. coli* respectively at the low temperature. Note that $r_P^L > r_E^L$ and the difference is defined as $\Delta r^L \equiv r_P^L - r_E^L$. Similarly, r_P^H and r_E^H denote the specific growth rates of *P. putida* and *E. coli* respectively at the high temperature. $r_P^H < r_E^H$ and the difference is defined as $\Delta r^H \equiv r_E^H - r_P^H$. The above equation shows that only the ratio of the two time intervals matters, and it is equal to the inverse of the ratio of specific growth rate differences between the two species at the corresponding temperatures. This is a specific quantitative relationship that only holds for the exponential growth model. However, certain qualitative aspects of this relationship are generalizable. In particular, it can be expected that the time interval at which there is a smaller grow rate difference between the two species needs to be longer.

In addition to enabling co-existence, the cycling temperature regime also provides an effective means for tuning the community composition, which is determined by the two time intervals and the initial composition. Taking the simple exponential growth model as an illustrative example again and using the ratio of the two species averaged over a cycle to quantify the composition, we can show that the average ratio over a cycle consisting of t^H at the high temperature and t^L at the low temperature is as follows (see Materials and Methods for details):

$$\left(\frac{E}{P}\right)_{\text{average}} = \left(\frac{E}{P}\right)_{\text{initial}} \cdot \left[\frac{\Delta r^L\left(e^{\Delta r^H \cdot t^H} - 1\right) + \Delta r^H\left(e^{\Delta r^L \cdot t^L} - 1\right)}{\Delta r^H \cdot \Delta r^L \cdot (t^H + t^L)}\right]$$

where $\left(\frac{E}{P}\right)_{\text{initial}}$ and $\left(\frac{E}{P}\right)_{\text{average}}$ represent the ratio of *E. coli* cells to *P. putida* cells at the beginning and averaged over the cycle, respectively. A desired average ratio, therefore, can be achieved when one sets the initial ratio and the two time intervals appropriately. In more general cases where the simple exponential growth model does not apply, the above well-structured analytical relationship may not exist. However, there will still be general trends. For instance, the shorter the two time intervals are, the closer the average ratio is to the initial one. Such design principles can be explored systematically through numerical simulation for specific systems.

Guided by the mathematical modeling and analysis described above, we experimentally implemented the cycling temperature regime on the *E. coli* and *P. putida* bi-culture model system. We were able to demonstrate that the bi-culture could be maintained over an extensive period of time when the parameter settings were chosen appropriately. A representative example is shown in Figure 2.4, where 37 ˚C was chosen as the high temperature favoring *E. coli* and 31 ˚C

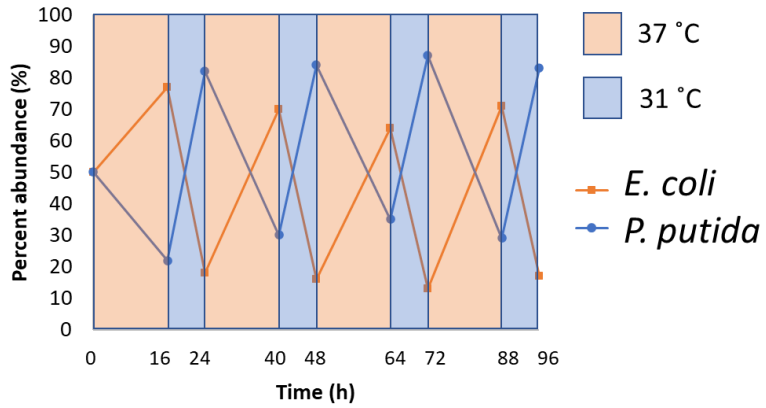


Figure 2.4 The cycling temperature regime is capable of maintaining coexistence of both species in the bi-culture. With alternating temperatures of which one favors E. coli and the other P. putida, each species is spared from competitive exclusion.

as the low temperature favoring *P. putida*. The bi-culture was started with an inoculation ratio of 1:1 and incubated in a cycling temperature program of 16 hours at 37 ˚C and 8 hours at 31 ˚C. It was maintained stably at an average composition of 45% *E. coli* and 55% *P. putida* with a standard deviation of 27% for a total of 4 cycles (96 hours).

As illustrated in Figure 2.4, the community composition exhibits dynamic oscillations under the cycling temperature regime. The amplitude of the oscillation, i.e. to what extent the composition fluctuates over time, is an important criterion that deserves closer examination. We investigated this first through computational simulation. Here, we assumed no direct

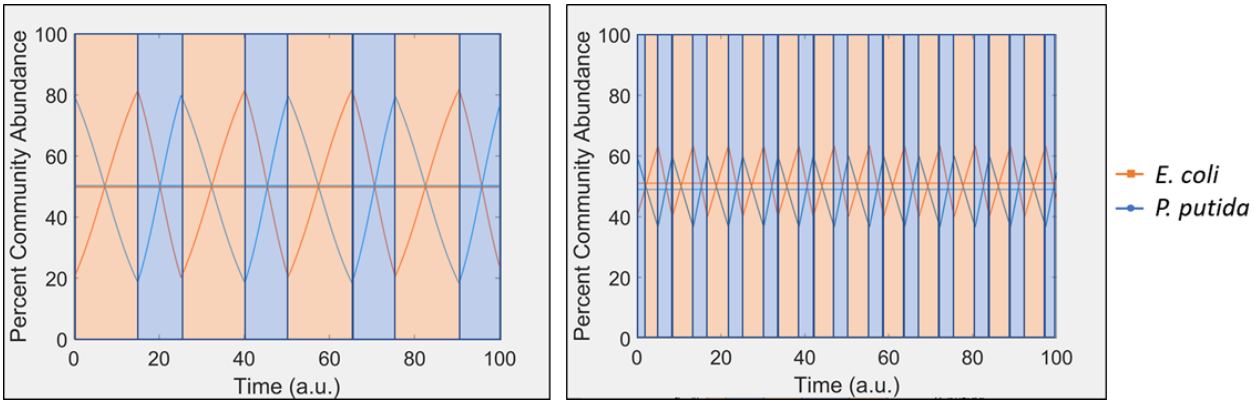


Figure 2.5 Effect of time intervals on the amplitude of community composition oscillation. Computational simulation shows that when growth related parameters are kept constant, longer time intervals lead to larger amplitude of oscillation (A: $t^H = 10.15, t^L = 15$), whereas shorter time intervals result in more desirable, smaller amplitude of oscillation (B: $t^H = 3.4, t^L = 5$). In each graph, the horizontal lines represent the average percentages, which are around 50%.

interactions between the two species and used two un-coupled logistic growth equations to model the bi-culture incubated in a cycling temperature program and diluted at the end of each time interval (see Materials and Methods for details). Our results showed that the amplitude of the oscillation was dependent on the duration of the time intervals. Longer time intervals result in larger amplitude oscillations in community composition (Figure 2.5 A) and shorter time intervals lead to more desirable, smaller amplitude oscillations (Figure 2.5 B). We next validated our model predictions with wet lab experiments. Figure 2.6 illustrated the results of an

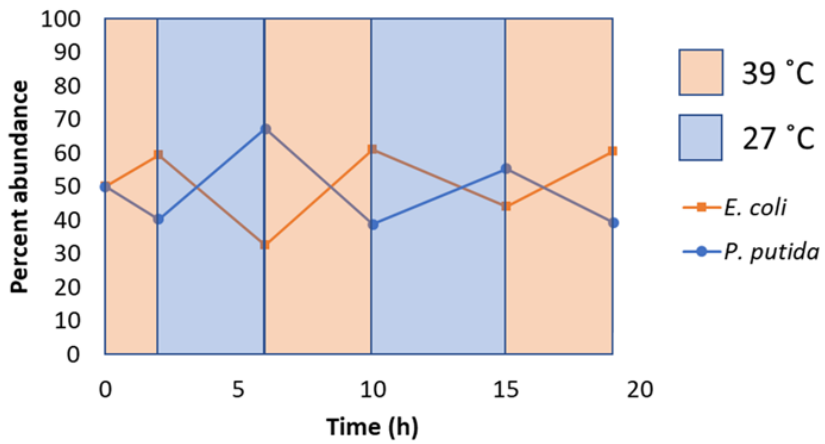


Figure 2.6 Experimental results demonstrated that shorter time intervals lead to oscillations of smaller amplitudes, in comparison to those in Figure 2.4.

experiment in which we set the temperatures to be 27 °C and 39 °C, which favor *P. putida* and *E. coli* respectively, and incubated the bi-culture with shorter time intervals compared to those in Figure 2.4. The bi-culture was maintained at an average composition of 51% *E. coli* and 49% *P. putida* with a standard deviation of 11%, which was substantially smaller than that in Figure 2.4, 27%. This approach of employing shorter time intervals can, theoretically, reduce the amplitude of oscillation in community composition to any arbitrarily small value. In practice, however, various constraints would arise. In particular, it takes cells a certain amount of time to physiologically respond to changes in temperature, which sets a lower limit on the time interval parameter in this cycling temperature regime.

2.4 Discussion

Synthetic microbial communities are increasingly being used for a variety of industrial, medical, and environmental applications such as production of commodity chemicals, proteins, and bioremediation (Höffner & Barton, 2014; S. Zhang, Merino, Okamoto, & Gedalanga, 2018). Because microbial community function depends in part on the community composition, it is of inherent interest to synthetic ecology, an emerging area of synthetic biology, that environmental conditions have been demonstrated to affect community composition in an

ecological context. While the synthetic biology/ecology literature has yet to treat the subject thoroughly, ecologists have been studying the effects of environmental factors on community composition for decades. It has been proposed and demonstrated that fluctuations in environmental conditions such as nutrient availability and temperature can enable coexistence and influence compositions of various natural communities (Harder & Veldkamp, 1971; Hutchinson, 1961; Jiang & Morin, 2007; Tilman, 1999). Interestingly, in the biotechnology context, the effect of modulating a variety of environmental parameters was explored in the 1970's and 1980's (Davison & Stephanopoulos, 1986). These previous studies, however, focused on the chemostat setting, which is not commonly used in biotechnological applications. It is known that an abundance of cellular processes such as nutrient sensing, signal transduction, and macromolecular synthesis are affected by the constantly changing environment within a batch culture (Ziv, Brandt, & Gresham, 2013). Therefore, the population dynamics in a batch culture are expected to differ significantly from those seen in a chemostat culture. Accordingly, here we show that in sequential batch cultures, the environmental condition of temperature can be used to modulate growth rates and rationally regulate community composition in synthetic microbial communities.

We show that constant temperature programs can be used to enforce competitive exclusion or to enable co-existence in a synthetic co-culture. It is worth noting that interestingly, the mean community composition does not respond symmetrically to changes in temperature (Figure 2.3 B). Specifically, we refer to the dramatic shift in community composition from 35.7 ˚C to 35.5 ˚C. Above 35.7 ˚C, the community composition gradually shifts towards higher abundance of *E. coli* as the temperature increases. In drastic contrast, the

community composition changes abruptly from primarily *E. coli* to entirely *P. putida* when the temperature decreases by a mere 0.2˚C from 35.7˚C to 35.5˚C. We do not understand the exact cause of this asymmetric behavior, but speculate that interspecies interactions play a role. It is possible that interactions commonly observed between species in microbial co-cultures such as nutrient competition (Ghoul & Mitri, 2016), secretion of toxins (Ghoul & Mitri, 2016), and interspecies crosstalk between quorum sensing systems (Ghoul & Mitri, 2016) occur in our system and could have effects on community composition. It would be interesting for future work to investigate the underlying mechanisms of this intriguing observation.

Control of community composition of synthetic microbial communities is important for optimizing functionality in a biotechnological context. We demonstrate here that temperature regulation is one modality that can be used to rationally regulate composition of synthetic microbial communities. However, a variety of other modalities such as pH, salt concentration, etc. can theoretically be implemented similarly, assuming that the basic requirements we propose for temperature are met (namely that one range of conditions favors growth of one species, and a second range of conditions favors growth of the other species). The advantage of temperature over some of these other modalities is that it can be dynamically changed without requiring media replacement and is readily implementable. Other modalities for regulating synthetic microbial community composition have been explored in the literature, most notably by Spencer Scott et al. (Scott et al., 2017). In this innovative work, the authors designed and implemented a self-limiting synthetic quorum sensing regulated lysis system to prevent competitive exclusion and enable coexistence between otherwise incompatible community members. One advantage of their approach is the potential for scalability in

community complexity, as there are a variety of orthogonal, well-characterized quorum sensing systems that could be utilized with different species or groups of species in a community. Whereas, the alternative approach demonstrated in this work has the advantage that because response to temperature is one of the oldest, core physiological responses of microbes, although adaptation to temperature is expected to diminish its effectiveness at some point, the time scales at which temperature regulation is able to enable coexistence and regulate community composition are quite long (up to 400 hours with our experimental system). However, it would not be straightforward to extend the cycling temperature regime to more complex communities; development of significantly more complicated control schemes will be required.

In this work we have explored temperature as a modality for rationally regulating community composition of a synthetic microbial consortium. While we do not expect the empirical values such as the specific temperatures employed to be applicable to different systems, the conceptual framework along with the associated design principles we have established can be applied to a wide variety of synthetic co-cultures. New approaches for regulating synthetic microbial community composition will continue to emerge and we envision that temperature based control schemes will contribute to a powerful toolbox in the future consisting of various well-developed modalities for regulating synthetic and natural microbial communities, from which researchers can choose based on the specific properties and performance requirements of their system.

Chapter 3: Distributed Sensing with Centralized Memory Utilizing Bacteriophage-mediated Information Transfer in Synthetic Microbial Communities

3.1 Introduction

The ability to extract information from the external or internal environment and transform that information into a specific biological response is vital to the survival of all living cells. Consequently, biological systems have evolved a multitude of solutions to this challenge over millions of years. More recently these systems have garnered interest from the biotechnology and synthetic biology communities for their potential to coordinate specific biological activities and processes with specific external and internal conditions, which is a valuable capability within a technological and engineering context. Systems which are capable of this feat, extracting some type of information from the environment and transforming that information into a biological signal or process, are often broadly referred to as biosensors and will be referred to as such hereafter.

3.1.1 Biosensors: extracting information from the environment

One of the most common types of biosensors is the inducible transcription factor. Transcription factors are a class of proteins that bind specific genetic sequences referred to as operators and help recruit additional factors (such as RNA polymerase) to operator adjacent

promoters to initiate transcription of downstream genes. Sometimes, in addition to transcription factor binding sites, there are repressor binding sites adjacent to the promoter, in which specific proteins referred to as repressors will bind and physically block the transcription machinery, thus preventing transcription initiation. In this way, transcription factors (and repressors) are able to turn gene expression on and off. Additionally, many transcription factors and repressors themselves have evolved the capability to be turned on and off in response to specific environmental stimuli. For example, a ligand-inducible transcription factor will have a specific physical conformation in the "off" state, and in that state it is physically incapable of binding its operator to recruit transcription initiation factors, and gene expression is therefore turned off. In the presence of its ligand, the transcription factor will undergo a conformational change after which it is then capable of binding its operator site and recruiting transcription initiation machinery thereby turning on gene expression from that site. Inducible transcription factors have evolved for a wide variety of ligands including metabolites, metals, etc. (Mahr & Frunzke, 2016). In addition to ligand-inducible transcription factors, there are transcription initiation systems that respond to light, temperature, pH, etc. (Mahr & Frunzke, 2016). In each case, the inducible transcription factor system functions as a biosensor by extracting some type of information from the environment (i.e. presence or absence of light, temperature, presence and sometimes even concentration of a ligand) and transforming that information into a biological process; in this case, expression of a specific gene. For many of these systems, it is possible to place nearly any gene of interest downstream of the inducible transcription system, which enables the coupling of expression of any gene of interest to any environmental signal for which there exists a biosensor. Given their obvious potential

environmental, industrial, and medical applications, there is an entire subfield of synthetic biology that has evolved around identifying, developing, and improving biosensors (Juárez, Lecube-Azpeitia, Brown, Johnston, & Church, 2018).

3.1.2 Biomemory: stable recording of information across generations

Perhaps even more important and fundamental to the survival of biological systems than the ability to sense and respond to a wide variety of environmental stimuli is the ability to stably record, copy, and modify information. This capability can be interchangeably referred to as biomemory. One could even argue that at their core, biological systems are simply packaged repositories for informational content accompanied by instructions and machinery to copy and make stochastic, intermittent modifications to that information such that natural selection can occur on these informational variants to select information packages which are most successful at replication and survival. In any event, it is inarguable that the ability to store and transmit information is fundamental to biological systems. The most common and well-known biological method for storing information is, of course, deoxyribonucleic acid (DNA). The modern era of molecular biology was ushered in by the ability to make targeted modifications to DNA. More recently, the discovery of CRISPR-Cas systems, which further refined that ability, has revolutionized the fields of molecular biology, synthetic biology, and biotechnology, among others. While CRISPR initially gained fame for the ability to (more or less) precisely cut DNA with easily modifiable specificity, that is really only one half of the story of CRISPR. The CRISPR system originally evolved as a form of microbial immunological memory (Barrangou, 2015; Nuñez, Lee, Engelman, & Doudna, 2015).

In this immunological memory system, double stranded DNA breaks are first bound by the recBCD complex, a very well characterized molecular machine involved in the canonical DNA damage repair pathway (Jon McGinn & Luciano A. Marraffini, 2019; Smith, 2012). In an interaction that is currently not completely understood, the recBCD complex performs some type of preprocessing of the double stranded DNA template to produce single stranded DNA substrate which then becomes available for interaction with the CRISPR adaptation Cas1-Cas2 complex (Figure 3.1 top right) (Jon McGinn & Luciano A. Marraffini, 2019). The Cas1-Cas2 complex then excises approximately 33 base pair fragments of the single stranded DNA and inserts those fragments into the CRISPR array as new spacers. When expressed at its native concentrations, the Cas1-Cas2 complex exhibits specificity for foreign DNA sequences, but the molecular basis for this specificity is currently unknown. It is tempting to hypothesize that some type of spatial sequestration of the Cas1-Cas2 complex at the membrane and away from genome replication machinery could contribute to this specificity. What is known is that when the Cas1-Cas2 machinery is overexpressed it loses its specificity for non-native DNA and will target any DNA substrate.

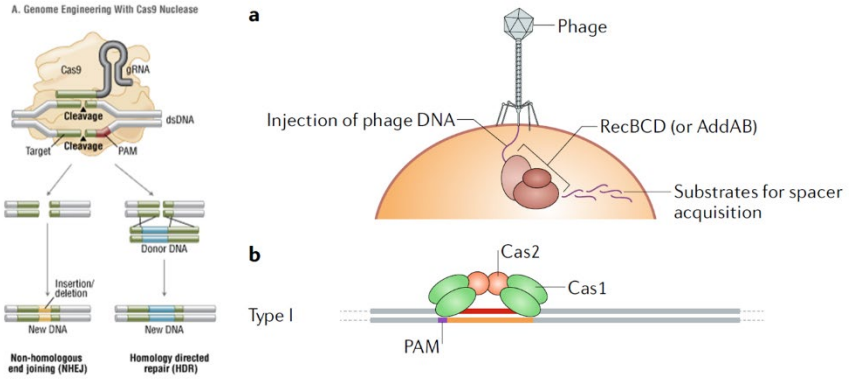


Figure 3.1 Left, a cartoon depicted CRISPR mediated endonuclease activity followed by homologous recombination based repair after cleavage (Reis, Hornblower, Robb, & Tzertzinis, 2014). Right, a) injection of phage DNA into a target cell followed by preprocessing of DNA by the recBCD complex b) excision of a DNA fragment that will become a new spacer in the CRISPR array by the Cas1-Cas2 complex as part of the CRISPR adaptation response (J. McGinn & L. A. Marraffini, 2019)

In any event, under naturally occurring conditions, the DNA sequences introduced into the CRISPR array as new spacers therefore serve as a form of molecular memory of the sequences of foreign DNA to which the cell has been previously exposed. This so-called CRISPR adaptation response, then, provides a form of easily modifiable biological memory that can be used as a repository for an extremely wide array of informational content (4 possible base pairs at each of 33 positions gives ~7.38×10^{19} possible unique sequences for a single spacer, and arrays are composed of ~30-50 spacers on average depending on the CRISPR system class, type, and subtype) (Toms & Barrangou, 2017). Importantly, new fragments are always added to the same end of the CRISPR locus, which enables the ability to encode chronological information about the order in which DNA sequences were encountered into the order in which they appear in the CRISPR array.

3.1.3 Cell-to-cell information transfer in biological systems

3.1.3.1 Introduction to information transfer in biological systems

Having established the means to sense information from the environment, encode that information into biologically compatible molecules capable of inducing a specific response or process, and when appropriate stably recording that information, the final biological process that is pertinent to our discussion is the ability of cells within a community to collectively share and process information when necessary. Within a single population of a single species/cell type, a variety of collective behaviors are known to be regulated by shared information in bacterial communities such as bioluminescence, the secretion of virulence factors, production of public goods, and the formation of biofilms (Miller & Bassler, 2001; Papenfort & Bassler, 2016). Arguably the most well-characterized microbial information sharing system is commonly

referred to as quorum sensing. In quorum sensing systems, each cell within the population secretes a low concentration of a signaling molecule (Miller & Bassler, 2001; Papenfort & Bassler, 2016). At low population density of that cell type, because the concentration of signaling molecule secreted by each cell is low, signal is unable to accumulate (Miller & Bassler, 2001; Papenfort & Bassler, 2016). At high population density, however, signal is able to accumulate above a threshold which activates quorum sensing receptors that specifically bind and respond to the quorum sensing molecule (Miller & Bassler, 2001; Papenfort & Bassler, 2016). These receptors then initiate a biological response to this population density information, often in the form of gene regulation (Miller & Bassler, 2001; Papenfort & Bassler, 2016). Evolutionarily speaking, it is understandable that cells would want to share information regarding population density before upregulating genes related to collective behavior such as colonization and invasion. Although quorum sensing is one of the best known and well-characterized information sharing and processing systems in bacteria, there are a variety of other mechanisms for information transfer in microbial communities such as vesicle-mediated horizontal gene transfer, conjugative transposon transfer, bacteriophage-mediated gene transfer, etc. (Chiang, Penadés, & Chen, 2019; Guynet, Le, Chandler, & Ton-Hoang, 2020; Tran & Boedicker, 2019). However, the ways in which biosensing components, biomemory components, and information sharing and processing systems are linked into networks of multiple species or cell types in microbial communities is relatively poorly studied and characterized.

3.1.3.2 Bacteriophage lambda biology within an information transfer context

As mentioned above, bacteriophage- mediated gene transfer (also referred to as transduction) is another way that information can be transmitted from one cell to another within microbial populations. Bacteriophage are a diverse group of viruses that replicate within prokaryotic organisms (Hendrix, 1983). It is extremely well documented that bacteriophage-mediated gene transfer plays an important role in horizontal gene transfer in many microbial communities (Eggers et al., 2016; Keen et al., 2017; Sutton & Hill, 2019; Valero-Rello, López-Sanz, Quevedo-Olmos, Sorokin, & Ayora, 2017; von Wintersdorff et al., 2016). Bacteriophage lambda is a specific bacteriophage of the family *Siphoviridae* genus *Lambdavirus* species *Escherichia virus Lambda* that replicates in *E. coli* cells and is one of the best studied model systems in molecular biology (ExPASy; Hendrix, 1983). The availability of a wealth of information regarding the genetics and biochemistry of bacteriophage lambda combined with its ability to manufacture and package a large (~50 kb) genetic payload of double-stranded DNA make bacteriophage lambda an excellent candidate for a mechanism to transfer information in *E. coli* populations.

3.1.3.2.1 Bacteriophage lambda virion morphology and packaging

Lambda phage particles (also referred to as virions) consist of approximately 50% protein and 50% DNA (Hendrix, 1983). Each particle contains a single double-stranded DNA molecule encapsulated in the capsid head, which is an icosahedral structure 0.05 microns in diameter attached to a tubular tail which is approximately 0.15 microns long (Figure 3.2) (Casjens & Hendrix, 2015; Hendrix, 1983). The native lambda genome is ~48 kb but DNA molecules ranging from ~75-105% of that size are competent to be packaged into a lambda capsid head (Hershey; B. Hohn & Murray, 1977). In the mid-1970s and early 1980s the DNA recognition sequences that are necessary and sufficient for packaging of DNA into the capsid head were discovered and characterized by Barbara Hohn, enabling packaging of heterologous DNA sequences into lambda phage capsid heads (Barbara Hohn, 1975; B. Hohn, 1983; B. Hohn & Murray, 1977). Lambda phage particles are metabolically inert in isolation but can adsorb to a target cell via virion tail fibers which bind to receptors on the host cell (Guan, Ibarra, & Zeng, 2019). The primary *E. coli* receptor for lambda phage is the lamB receptor, an outer membrane porin involved in maltose uptake, although mutations in the tail fiber proteins can also allow the lambda virion to interact with the ompC receptor as well (Guan et al., 2019; Hendrix, 1983). Upon adsorption to a target cell (also referred to as host cell) the DNA payload is injected via

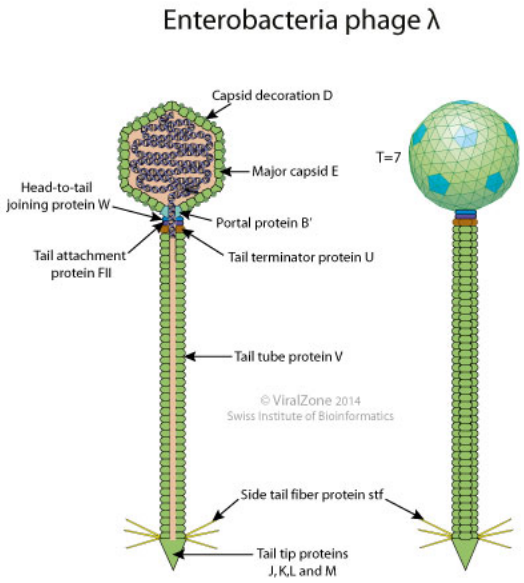


Figure 3.2 Cartoon depiction of bacteriophage lambda particle head and tail morphology (ExPASy).

the lambda phage tail (Inamdar, Gelbart, & Phillips, 2006). While the forces responsible for this injection event have not been completely characterized, it is proposed that a significant driving force results from the high degree of stress subjected upon the viral DNA by its packaging into the viral capsid, and that this driving force is aided by ratcheting and entropic forces associated with proteins that bind to the viral DNA within the host cell cytoplasm (Inamdar et al., 2006).

3.1.3.2.2 Bacteriophage lambda life cycles

Upon injection, expression of lambda phage genes results in entry of the cell into the bacteriophage lysogenic or lytic life cycles (Shao, Trinh, & Zeng, 2019). The factors influencing the lysis-lysogeny decision in bacteriophage lambda are complex and have been the subject of extensive study but are outside the scope of this discussion (Shao et al., 2019). In the lysogenic pathway, the lambda phage DNA is integrated into the host genome and replicated along with the host genome, resulting in stable transmission from mother to daughter cells (Campbell, 2003). When incorporated into the host genome in this way, the phage genome is referred to as the prophage, and the phage genome can be maintained by the host cell in this latent prophage state indefinitely (Campbell, 2003). In the lytic life cycle, the biosynthetic machinery of the host cell is subverted from their normal tasks in cellular growth and division and reprogrammed to become phage production machinery. Copies of the phage genome are produced along with new phage particles, and the phage genomes are packaged into the newly synthesized phage virions (Campbell, 2003). Interestingly, and importantly for our purposes, cells in the lysogenic cycle harboring a prophage can be induced into the lytic cycle by a variety of cellular stress signals. The core signal for induction of lysogenic cells into the lytic pathway, however, is DNA damage.

3.1.3.2.3 Induction of lysogenic cells into the lytic cycle

The canonical signal for induction of lambda lysogens into the lytic cycle is DNA damage, mediated by the recA protein. The recA protein plays a central role in the DNA damage and response pathways in *E. coli* (Lusetti & Cox, 2002). In the absence of symptoms of DNA damage, recA exists in its inactivate state. In the presence of single-stranded DNA (ssDNA) (which is a common biomarker of DNA damage in the cell) and ATP, recA will polymerize onto the ssDNA and become activated as a protease (Lusetti & Cox, 2002). In its activated state (sometimes referred to as recA*), recA will catalyze autocleavage of the lexA repressor, which is a global regulator of the SOS response to DNA damage in *E. coli* (Butala, Zgur-Bertok, & Busby, 2009; Galkin et al., 2009). The lambda phage repressor protein cI is a structural homologue of lexA and also undergoes recA catalyzed self-cleavage when the SOS response is induced (Galkin et al., 2009). There are constitutively active mutants of recA (specifically recA730) which are sufficient to induce lambda lysogens into the lytic cycle when expressed intracellularly (McCall, Witkin, Kogoma, & Roegner-Maniscalco, 1987). Controlled expression of a constitutively active recA mutant therefore exhibits promise as the basis of a biosensor component capable of inducing the lytic response in lambda lysogens.

3.1.4 Overall approach

The goal of this project is to develop a biological system that is capable of combining biosensing elements with biomemory elements to create a biological system that is theoretically capable of sensing a large number environmental signals and recording the nature or identity of those signals in addition to the order of exposure to those signals in a microbial community. The capability to sense and record environmental information in a biological

system has been demonstrated in the past using a variety of approaches, but our system uses a novel distributed sensing architecture combined with a centralized memory component which spatially segregates the sensing components from the memory components into separate cellular population within a microbial community. This distributed architecture theoretically enables multiplexing of biosensing components to levels not achievable with previous systems while preserving the ability to record information about not only the identity of signals to which

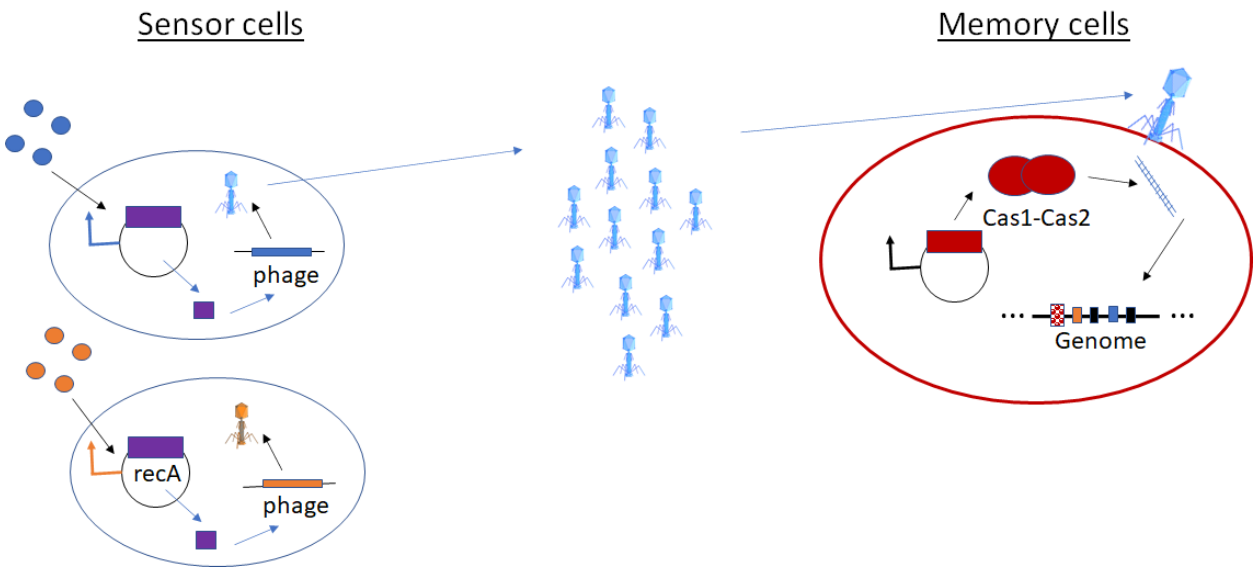


Figure 3.3 Cartoon depiction of system architecture with distributed sensing and centralized memory components enabled by bacteriophage-mediated information transfer. Sensor cells are depicted by large blue ovals, Inducer molecules for the biosensor component are depicted by small blue and orange extracellular circles, and the recA730 protein is depicted as a purple square. In the memory cells the Cas1-Cas2 complex is depicted by overlapping red ovals.

the community is exposed, but also the order in which the community was exposed to those signals. Because this separation between sensing and memory components requires transmission of information from the sensing components to the memory components (which has not previously been required, as both sensing and memory components were always contained within a single cell), we pioneer the use of a bacteriophage-based information transfer system to connect the distributed sensing components with the centralized memory component.

In brief, we designed a plasmid-based genetic circuit in *E. coli* that uses a constitutively active recA mutant (recA730) under control of an inducible promoter to activate lysogenic lambda phage in response to the inducing signal for the promoter. Upon entering the lytic cycle, these cells manufacture approximately 100 new virions per activated cell, after which the cells are lysed, releasing the virions (Greene & Rao, 1998). The sensors cells are co-cultured with memory cells, which are a strain of *E. coli* overexpressing the CRISPR adaptation complex Cas1-Cas2, which is necessary and sufficient for excising and processing 33 bp fragments from intracellular DNA substrate and subsequently integrating these fragments into the CRISPR array as new spacers. Upon release of virions from the sensor cells, the virions adsorb to memory cells and inject their lambda DNA cargo which is then targeted as substrate by the CRISPR adaptation machinery, resulting in the addition of new spacers into the CRISPR array that map back to the lambda genome (Figure 3.3). In addition to new spacers that map back to the lambda genome, the overexpressed Cas1-Cas2 complex also attacks the *E. coli* genome itself (as overexpression of Cas1-Cas2 leads to loss of specificity for foreign DNA fragments), and therefore at a population level, a mix between new spacers that map back to the lambda genome and new spacers that map back to the *E. coli* genome is observed.

3.2 Materials and Methods

3.2.1 Microbial strains and cultivation

E. coli K12 MG1655 was used as a cloning chassis for plasmids unless otherwise indicated. *E. coli K124* was used as the lysogenic host strain for lambda prophage unless otherwise indicated. Cultures were grown in sterile autoclaved Fisher Scientific Lysogeny Broth- Lennox (LB-Lennox) or minimal M9 media (200 mL 5x M9 salts [34 g/L Na_2HPO_4

anhydrous, 15 g/L KH_2PO_4, 2.5 g/L NaCl, 5 g/L NH_4Cl], 20 mL 20% glucose, 2 mL 1M $MgSO_4$, 100 uL 1M $CaCl_2$, 780 mL sterile deionized H_2O; autoclave all components separately then mix under sterile conditions) as indicated.

3.2.2 Cryostocks

Unless otherwise indicated, cryostocks were generated by mixing 500 uL saturated overnight culture with 500 uL sterile glycerol and then stored at -80 ˚C.

3.2.3 Antibiotics and inducers

Unless otherwise indicated, antibiotic and inducer concentrations were as follows: Chloramphenicol 35 ug/mL; Kanamycin 50 ug/mL; anhydrotetracycline (aTc) 100 ng/mL (aTc induces the pRec plasmid encoding the Cas1-Cas2 complex in the memory strain)

3.2.4 Miniprep

Minipreps were performed using the Qiagen Qiaprep Spin Miniprep Kit.

3.2.5 Polymerase Chain Reactions (PCR)

Unless otherwise indicated, PCR reactions were performed using NEB Phusion High Fidelity DNA Polymerase (M0530) with the thermocycling conditions indicated by NEB and annealing temperatures suggested for each primer pair by the ThermoFisher Scientific annealing calculator (ThermoFisherScientific).

3.2.6 recA constitutively active mutant generation

E. coli K12 MG1655 was cultured overnight in LB-Lennox at 37 ˚C with 225 rpm shaking. The genome was then purified using the Qiagen DNeasy Blood and Tissue kit. Wild type recA

was PCR amplified from the *E. coli K12 MG1655* genome as discussed in Materials and Methods section 3.2.5 using primers:

recA_WT_SacI_for : acagcgGAGCTCatggctatcgacgaaaacaaaca and

recA_WT_HindIII_rev : acagcgAAGCTTttaaaaatcttcgttagtttctg

This amplicon and the plasmid pB33eCPX (Rice & Daugherty, 2008) were digested with SacI and HindIII, ligated together with T4 DNA Ligase (NEB), electroporated into *E. coli K12 MG1655* and plated on LB + Chloramphenicol plates. A single colony was isolated and cultured overnight in LB-Lennox + Chloramphenicol and cryostocked. An overnight culture was grown from this cryostock in LB-Lennox + Chloramphenicol and miniprepped. The single base pair mutation in recA to generated recA730 was generated with PCR using the primers:

recAWT_E38K_for: accgttccatggatgtgAAAaccatctctaccggttcgc and

recAWT_E38K_rev: cttcacccaggcgcatgatggag

This amplicon digested with the restriction enzyme DpnI (NEB) to eliminate template plasmid, 5' phosphorylated with T4 Polynucleotide Kinase (NEB), electroporated into *E. coli K12 MG1655* and plated on LB+Chloramphenicol plates. A single colony was isolated and cultured overnight in LB-Lennox + Chloramphenicol and then cryostocked and miniprepped. Mutagenesis was confirmed via Sanger sequencing with the sequencing primer:

pB33recA_WT seq for: ccagatgggcattaaacgagtat

A ribosome binding site (RBS) for recA730 was added using the PCR primers:

recA_RBS_for: AGGAGTAAAAatggctatcgacgaaaacaaacagaaagc

recA_RBS_rev: TTTTACTCCTGagctcgaattcgctagcccaaaaaaac

This amplicon was digested with the restriction enzyme DpnI (NEB) to eliminate template plasmid, 5' phosphorylated with T4 Polynucleotide Kinase (NEB), electroporated into *E. coli K12 MG1655* and plated on LB + Chloramphenicol plates. A single colony was isolated and cultured overnight in LB-Lennox + Chloramphenicol and then cryostocked and miniprepped. Mutagenesis was confirmed via Sanger sequencing with the sequencing primer:

pB33recA_WT seq for: ccagatgggcattaaacgagtat

to generate plasmid pB33recA730RBS.

3.2.7 recA730 phenotype characterization

Plasmid pB33recA730RBS was then electroporated into *E. coli* K124 lambda lysogen and plated on LB + Chloramphenicol plates. A single colony was isolated and cultured overnight in LB-Lennox + chloramphenicol and cryostocked. *E. coli* K12-pB33recA730 was then seeded from the -80 ˚C cryostock into LB + Chloramphenicol and cultured overnight at 37 ˚C with shaking at 225 rpm. The next day the saturated overnight culture was sub-cultured into fresh LB + Chloramphenicol media and grown to OD600 = 0.4 at 37 ˚C with 225 rpm shaking. At this point, the culture was split into two conditions, a negative control and the experimental condition; for each condition 2 mL of culture was added to a 14 mL round-bottom Falcon tube. 1% arabinose was added to the experimental condition and an equal quantity of sterile de-ionized water was added to the negative control. The cultures were then cultured for 2.5 hours.

3.2.8 Lambda virion isolation

Lambda lysogens were overnighted in LB-Lennox + chloramphenicol at 37 ˚C with 225 rpm shaking. This overnight was passaged into fresh LB-Lennox + chloramphenicol media and grown to OD600 = 0.4. Cultures were then induced with either 1% arabinose (if carrying pB33recA730) or 2 ug/mL mitomycin C and cultured for 2-3 hours at 37 ˚C with 225 rpm shaking until culture clearing was observed. Cultures were then centrifuged at 3,000 x g for 3 minutes at room temperature in a tabletop centrifuge. The supernatant was then filtered with a 0.22 µm filter and stored at 4 ˚C protected from light.

3.2.9 Sensor strain co-culture with memory strain

Sensor strain *E. coli* K124-pB33recA730RBS and memory strain *E. coli* BL21(DE3)-pRec were cultured overnight in LB-Lennox + chloramphenicol at 37 ˚C with 225 rpm shaking. Overnights were sub-cultured into fresh LB-Lennox + chloramphenicol the next day and grown to OD600 ~ 0.4; the cell densities were then normalized to each other using their OD600 reading. Sensor and memory cells were inoculated 1:250 in fresh LB-Lennox + chloramphenicol + aTc + 0.2% arabinose + 0.2% maltose + 10 mM MgSO4 or a no arabinose control. The ratio of sensor to memory cells was 1:100. Cultures were incubated overnight at 37 ˚C with 225 rpm shaking. The next day 800 uL of culture was spun down at 3,000 x g for 3 minutes, the supernatant was discarded and the cells were resuspended in 500 uL deionized water, and boiled for 15 minutes at 95 ˚C.

3.2.10 PCR amplification of newly expanded CRISPR arrays

0.2 uL of the boiled lysate from the sensor strain co-culture with memory strain is added to a 50 uL PCR reaction with Phusion (NEB) polymerase. The primers used were:

oYZ696: GTTGGTAGATTGTGACTGGC

oYZ694-MG7F: ATTTTGCGTTTCGTTCAGGT

The amplicon was run out on a 1.5% agarose gel and the band at ~420 bp was gel extracted using the QIAquick Gel Extraction Kit and eluted in 20 uL EB.

3.2.11 Sequencing of expanded CRISPR array amplicons

Samples of CRISPR expanded array amplicons were submitted to the Massachusetts General Hospital CRISPR sequencing core and results were analyzed with the Integrative Genomics Viewer (IGV) computer program.

3.2.12 Cosmid generation

E. coli K124 was inoculated from cryostock and cultured overnight at 37 ˚C with 225 rpm shaking. The next day the genome was extracted using the Qiagen DNeasy Blood and Tissue kit. The lambda packaging sequence was PCR amplified from the *E. coli* K124 genome with the primers:

Fragment.FOR: ccaaatactgtccttctagtgtagccgtagttaggccacCatacgatacctgcgtcataattgattatttgac

Fragment.REV: gctacagagttcttgaagtTTAATTAAcgctgtacagcgCacttccattgttcattccacggac

Plasmid pB33rec730RBS was amplified with primers:

pB33recA730RBS_packadd_for: acagcgTTAATTAAacttcaagaactctgtagca

pB33recA730RBS_packadd_rev: acagcgCTCGAGgtggcctaactacggctaca

These amplicons were combined in a NEBuilding HiFi DNA Assembly reaction and transformed into NEB® 5-alpha Competent *E. coli* (Subcloning Efficiency) and plated onto LB + chloramphenicol. Single colonies were isolated and cultured overnight in LB-Lennox + chloramphenicol and miniprepped. The resulting sample was submitted for Sanger Sequencing using primer:

lambda_chi_check_rev: gtaacctgtcggatcaccgg

This plasmid was then transformed into *E. coli* K124 and co-cultured with a memory strain, the expanded CRISPR array was amplified and submitted for sequencing as described above.

3.2.13 Lambda phage infection with isolated virions

Indicator cells are inoculated from cryostock and cultured overnight at 37 ˚C in tryptone broth (1% Bacto-tryptone; 0.5% NaCl; mixed in deionized water and autoclaved). The next day cultures are centrifuged at 3000 x g for 3 minutes at room temperature and the supernatant is discarded. The cell pellet is resuspended in ½ volume TMG (10 mM Tris-HCl pH 8; 10 mM MgSO4; 10 ug/mL gelatin) and incubated with shaking at 37 ˚C for 1 hour with 225 rpm shaking. Cultures are then pipetted into 100 uL aliquots in 14 mL round bottom Falcon tubes and chilled on ice. 50 uL of isolated phage virion is added to the aliquots and mixed gently. The mixture is let stand at room temperature for 15-20 minutes and then mixed with 3 mL TOP agar (tryptone broth + 7.5 g/L agar; autoclaved and maintained molten at 52 ˚C in a water bath) and immediately plated onto a tryptone plate (tryptone broth + 15 g/L agar) and incubated overnight at 37 ˚C.

3.2.14 Recombinant lambda phage genomes

Recombinant lambda genomes were generated as follows. For insertion of chi sites, the kanamycin cassette from pSAS31 (Scott A. Scholz et al., 2019) was amplified and inserted into pB33recA730RBS with NEBuilder HiFi DNA assembly. Four tandem chi sites were then inserted into the plasmid adjacent to the kanamycin cassette. The kanamycin cassette + chi sites were amplified using primers that included 40 bp homology domains for the lambda phage genome. This amplicon was electroporated into cells which had been prepared for lambda red recombineering according to the Court lab protocol (Sharan, Thomason, Kuznetsov, & Court, 2009) and plated onto LB + chloramphenicol + kanamycin plates. Single colonies were isolated and cultured overnight in LB + kanamycin at 37 ˚C and then induced for lambda virion isolation as described above. Lambda virions were used to infect naïve *E. coli K12 MG1655* cells as described above. All TOP agar and plaques generated from this infection were scraped into 14 mL round bottomed Falcon tubes with 5 mL LB-Lennox + kanamycin and cultured overnight. This liquid culture was then plated on Kanamycin plates and restreaked to isolate single kanamycin resistant colonies. For insertion of mNeonGreen, the mNeonGreen + kanamycin cassette was PCR amplified from pSAS31 (Scott A. Scholz et al., 2019) and electroporated into lambda red recombineering competent cells prepared according to the Court lab protocol (Sharan et al., 2009) and plated onto LB + chloramphenicol + kanamycin plates. Single colonies were isolated and cultured overnight in LB + kanamycin at 37 ˚C with 225 rpm shaking and then cryostocked.

3.2.15 Lambda red recombineering

Lambda red recombineering is a technique developed by the Court lab at MIT to insert or delete regions of the *E. coli* chromosome ranging from tens of base pairs to several kilobases (Sharan et al., 2009). Briefly, three genes from the lambda phage genome are expressed from a plasmid to facilitate homologous recombination of a desired DNA targeting construct into the *E. coli* genome (Sharan et al., 2009). The DNA targeting construct is electroporated into cells which are already expressing the lambda red genes from a plasmid and usually includes the gene cassette of interest and an antibiotic selection cassette flanked by 40-50 bp homology domains to the desired *E. coli* genomic site for integration of the targeting construct. The three lambda red proteins are the same three proteins used by lambda red phage to insert its own genome into the *E. coli* genome (Sharan et al., 2009). The Exo enzyme is a nuclease initiates degradation from 5' ends of linear dsDNA creating 3' ssDNA overhangs (Sharan et al., 2009). The beta protein is a ssDNA binding protein which then binds the ssDNA overhangs and protect them while also annealing them to complementary ssDNA in the cell (Sharan et al., 2009). Beta has also been reported to have strand invasion activity (Sharan et al., 2009). The gam protein inhibits the *E. coli* RecBCD and SbcCD exonuclease complexes to prevent degradation of linear ds DNA substrates (e.g. the targeting constructs) (Sharan et al., 2009). After electroporation of the targeting construct cells are recovered and plated on appropriate antibiotics (Sharan et al., 2009).

3.2.16 Recombinant lambda genome fluorescence characterization

Fluorescence evolution was characterized in the mNG recombinant lambda strain as follows. Cultures were inoculated from cryostock and incubated overnight at 37 ˚C in LB +

kanamycin with 225 rpm shaking. Saturated overnight cultures were then pipetted into 200 uL volumes in a 96 well plate and measured for fluorescence with Excitation= 485 nm and Emission= 515 nm. *E. coli* K124 cells with the chi site kanamycin cassette inserted via lambda red recombineering were used as a negative control.

3.3 Results

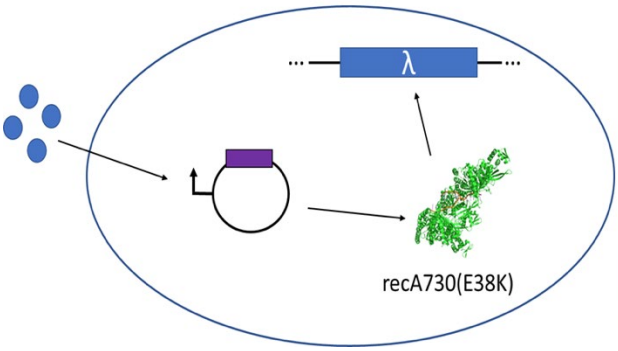
3.3.1 Biosensor activator for lysogenic lambda phage

For a biosensor component that is compatible with our overall design, we required a sensor that is capable of activating lysogenic lambda prophage in the sensor cells in response to a specific stimulus. To accomplish this, we started with the ligand-inducible promoter pBAD and the corresponding regulatory protein araC. In this commonly used promoter system, the inducing ligand is the sugar arabinose. In the absence of arabinose, the araC protein adopts a conformation that binds to its operator adjacent to the pBAD promoter and prevents transcription machinery from binding the promoter by inducing a DNA looping architecture (Martin & Rosner, 2001; Schleif, 2010). Upon binding to arabinose, the regulatory protein araC undergoes a conformational change that results in dissolution of the DNA looping structures and converts araC from a repressor into an activator protein, which actively recruits transcription machinery to the pBAD promoter, resulting in expression of any open reading frames downstream of the promoter (Martin & Rosner, 2001).

To connect the promoter with lambda phage induction, we utilized a constitutively active mutant of the recA protein, recA730 (McCall et al., 1987). Within a natural context, lysogenic phage maintains its latent state through activity of the phage repressor cI which

represses transcription/translation of the phage genes responsible for activating the lytic cycle (Hendrix, 1983). Cellular stress, specifically DNA damage, is a signal that will normally activate lysogenic prophage into the lytic cycle by converting the recA protein into its active conformation (Hendrix, 1983). Within the canonical DNA damage pathway, the protein recA is constitutively expressed but remains in an inactive conformation in the absence of DNA damage. Upon binding single-stranded DNA (ssDNA) and ATP (which are common substrate by-products generated by DNA damage), recA polymerizes and adopts its active conformation (commonly denoted as recA*). Upon activation, recA* interacts with the lambda phage repressor cI and catalyzes its auto-cleavage resulting in activation of lysogenic prophage into the lytic cycle (Galkin et al., 2009).

We therefore identified a constitutively active mutant of recA (recA730) and cloned this gene downstream of the arabinose inducible pBAD promoter in our plasmid (McCall et al., 1987). Conveniently, only a single base pair mutation (G112A) resulting in a single amino acid substitution (E38K) is required to convert wild-type recA into the constitutively active recA730 mutant (McCall et al., 1987). In Figure 3.4, a cartoon of *E. coli K12-MG1655* harboring a lambda prophage and a plasmid with our recA70 mutant under control of the pBAD promoter is shown. In Figure 3.4 A-B, a culture of these sensor cells was grown and



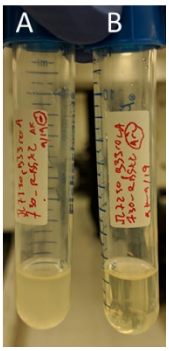


Figure 3.4 Left, a cartoon depiction of the sensor cells. The cell is represented by the large, blue oval. The inducing molecule arabinose is represented by the small, blue, extracellular circles. The black circle with promoter symbol represents the plasmid encoding the recA730 mutant under control of the pBAD promoter, and the lambda prophage is represented as a blue rectangle. Right, A) a culture of sensor cells without arabinose added and B) with arabinose added, resulting in phage production, cell lysis, and culture clearing.

then split into two conditions. In Figure 3.4 A, no arabinose is added and the culture continues to grow normally. In Figure 3.4 B, upon addition of arabinose the constitutively active recA mutant is expressed which results in induction of the lambda prophage, activation of lytic cycle, and consequent clearing of the culture after virion production and cellular lysis.

3.3.2 Memory component compatible with bacteriophage-mediated information transfer

For the memory component in our system we utilized the pRec plasmid developed by Sheth et al. 2017 and purchased from Addgene (Sheth, Yim, Wu, & Wang, 2017). This plasmid expresses the Cas1-Cas2 complex from a plasmid under tetR regulation. As the plasmid was initially developed to express Cas1-Cas2 for targeting plasmid DNA, we first sought to confirm that injection of lambda phage DNA into cells overexpressing Cas1-Cas2 would result in the addition of new spacers to the CRISPR array that map back to the lambda phage genome. We transformed the pRec plasmid into an *E. coli BL21* strain and cultured in the presence of tetracycline. We then introduced previously harvested and purified lambda phage virions (Figure 3.5). After 24 hours we PCR amplified the CRISPR locus in the BL21 cells and gel extracted the band corresponding to CRISPR loci which harbored a new spacer (the addition of a new spacer results in the insertion of a new 33 bp spacer in addition to replication of a 31 bp repeat, resulting in a 64 bp total addition to the PCR product which is easily resolvable with a 1.5% agarose gel) (Figure 3.5). The gel extracted PCR product was then submitted for

sequencing. In this initial experiment, we identified ~8% of new spacers that mapped back to the lambda genome (Figure 3.6) (the rest mapped back to the *E. coli* genome; in this system it is always expected that a portion of new spacers will map back to the *E. coli* genome as Cas1-Cas2 lose their specificity for foreign DNA when overexpressed and will target the *E. coli* genome in addition to foreign DNA fragments for spacer acquisition) (Sheth et al., 2017).

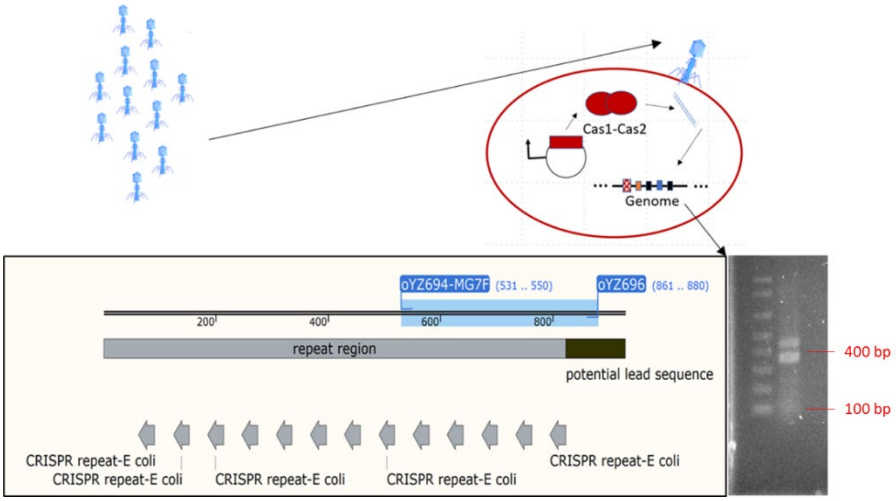


Figure 3.5 Top, a cartoon depiction of infection of the memory cells by lambda, followed by injection of lambda DNA and targeting of DNA by the Cas1-Cas2 CRISPR adaptation machinery for spacer excision and integration into the CRISPR array. Bottom left, a representation of the CRISPR array, including the region amplified by our primer set highlighted in blue. Bottom right, an agarose gel depicting the two product bands from a PCR reaction using an expanded CRISPR array as template. The product running at 350 bp is the amplicon from unexpanded CRISPR arrays and the production running around 420 bp is the amplicon from expanded CRISPR arrays with a new spacer.

After optimizing our infection conditions by including maltose in our culture media for higher expression of the lambda phage receptor lamB we detected ~25% of new spacers that mapped back to the lambda genome. Lastly, we directly co-cultured our sensor cells with the memory cells and induced the culture with arabinose. Under these conditions we identified ~50% of new spacers from the CRISPR array that mapped back to the lambda genome (Figure 3.6). While these experiments were preliminary and therefore only a single replicate was performed, these results provide evidence that the concept for our distributed sensing/centralized memory microbial community-based biosensor is viable. They also provide evidence that the genetic

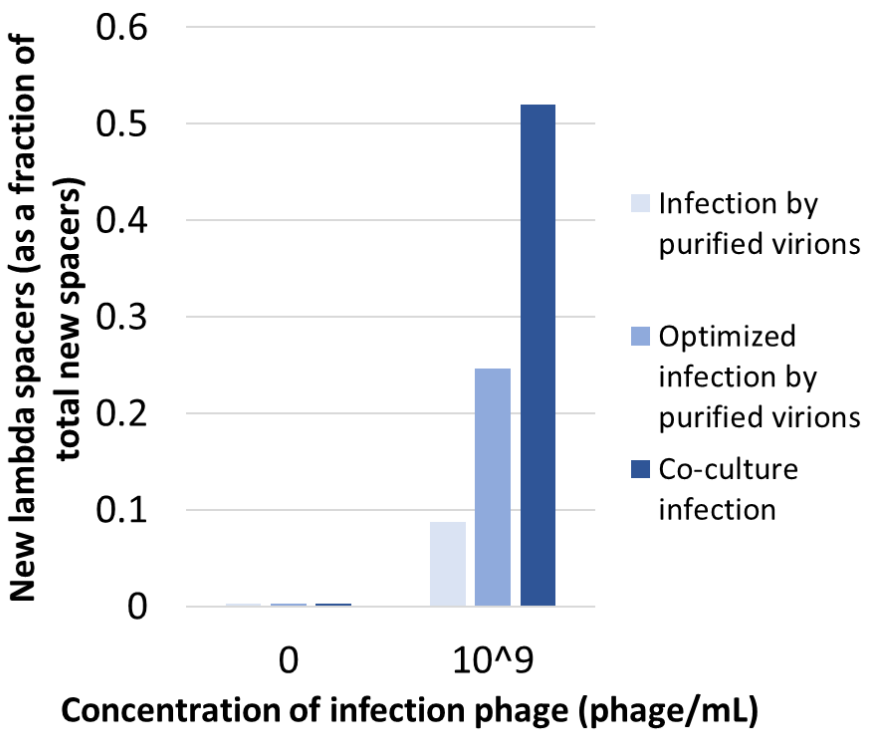


Figure 3.6 Bar graph depicting the fraction of new spacers identified in expanded CRISPR arrays that map back to the lambda genome. Results from three separate experiments, differentiated by color. Due to the preliminary nature of these experiments, only a single replicate was performed, therefore error bars are not possible.

circuit we developed for our biosensor component exhibits low/no activity in the absence of signal (i.e. arabinose), and upon introduction of signal the system is activated resulting eventually in introduction of new spacers into the CRISPR memory locus of the memory cells that map back to the lambda genome. By simply cloning the recA constitutively active mutant under the control of any biosensor components that control expression of a gene of interest, it should be possible to generate a pool of sensor cells that can be co-cultured with the memory population.

3.3.3 Encoding unique information into each memory event that correlates to the corresponding sensing event

3.3.3.1 Recording information correlated to sensing events as new spacers in CRISPR arrays

The final requirement to enable a biosensing/biomemory system with our architecture to record the identity and order of exposure of signals to which the community is exposed is that some type of unique identifying information which corresponds to the activating signal for

each sensor strain must be encoded into the information packaged into the virions produced by each different sensor strain population. Crucially, this unique identifying information must also be included in the new spacers added to the CRISPR array in memory cells during the adaptation response. If each biosensor activated an identical lambda phage signal, there would be no way to determine from which signals were introduced to the community and in what order by sequencing out the new spacers in the CRISPR array. We therefore began the process of developing a method to introduce this unique identifying information into the packaged information in each sensor strain by analyzing our system for opportunities.

We first investigated which sites in the lambda genome were selected as templates for excision and integration as spacers by the CRISPR adaptation response. Our idea was that if we could identify a region from which spacers were preferentially selected (so-called hotspots), we could clone unique DNA sequences into the particular region to impose uniqueness into the spacer acquired from each genome. Unfortunately, as can be seen in Figure 3.7, when the new spacer reads from the CRISPR locus are mapped back onto the lambda genome, the distribution and frequency of lambda genomic sites targeted as template for spacers appears to be entirely

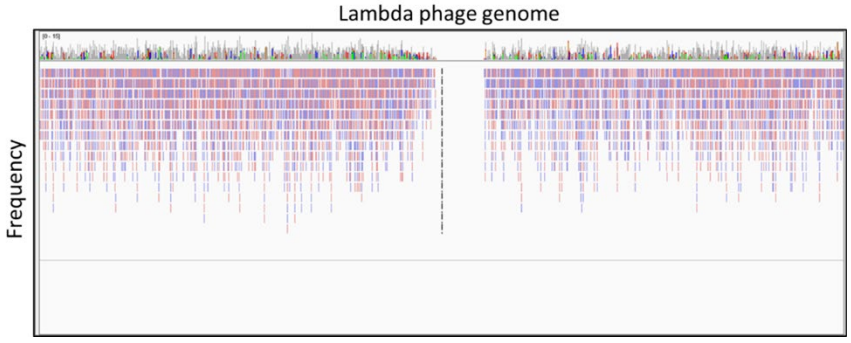


Figure 3.7 Frequency chart depicting the relative frequency with which sequences from different regions of the lambda phage genome are observed in new spacers of expanded CRISPR arrays. The lambda phage genome is spread across the X-axis and frequency is depicted along the Y-axis. The missing section of the lambda genome is due to replacement of a section of the B2 non-essential region of the phage genome with a kanamycin cassette (which was not incorporated into the reference genome). Colors and dotted lines to not depict relevant information.

stochastic. This feature of the system makes introducing uniqueness into the lambda genome challenging. We conceptualized two alternative approaches to overcome this obstacle.

3.3.3.1.1 Cosmid packaging

The first idea was that if we could package entirely different genetic information into the virions produced by each type of sensor cell, it would not matter what fragment of that DNA is targeted as template for new spacer acquisition because each signal would generate completely orthogonal sequences without homology. To implement this idea, we explored the use of cosmids, a technique developed in the late 1970s (Hohn 1977). Briefly, the DNA sequence that the lambda packaging machinery uses to identify substrate to be packaged into lambda virions was identified and characterized by Hohn et al. in the mid-1970s and early 1980s. The necessary and sufficient sequence was found to be approximately ~150 bp. Upon introduction of this sequence into any plasmid, the plasmid will subsequently undergo rolling circle replication until enough concatenated plasmid copies have been generated to fill a lambda virion head (~37.5-50 kb) and then be packaged. By introducing a different plasmid containing the lambda packaging sequence (without any homology to other plasmids used in the sensor pool) into each different sensor strain, it would be possible to package entirely orthogonal sequences into virions generated by different sensor cell types, and it therefore would not matter what region of that signal is targeted as spacer template by the CRISPR adaptation machinery in the memory cells after injection by the virions.

Subsequently we PCR amplified the lambda packaging sequence from the prophage in a *E. coli* lysogen and cloned it into the same expression plasmid we used for expression of our recA mutant (this first vector choice was for the sake of simplicity; in the future it could be

cloned into plasmids with no sequence homology to other plasmids used in different sensor cell types). We then co-cultured this sensor cell type with the memory population and sequenced the new spacer population that was PCR amplified from the memory cells. We found that very few new spacers mapped back to the lambda genome (0.2%), and approximately 2% of new spacers mapped back to the recA730 expression plasmid (30 reads mapping back to the lambda genome, 277 reads mapping to plasmid, and 10329 mapping to *E. coli* genome out of 10636 total reads). Due to the exploratory nature of this experiment, only a single replicate was performed. Additional replicates are needed to verify this result.

Assuming the result is valid, however, it is unclear whether the relatively small percentage of new spacer mapping back to plasmid DNA (compared to the proportion we usually find mapping back to the lambda genome) is due to inefficiency in packaging of plasmid DNA into virions or inefficient targeting of the injected DNA by the CRISPR adaptation response. The fact that we have observed a robust CRISPR adaptation response previously for phage injected DNA and that the same plasmid we used in our system has previously been used to generate Cas1-Cas2 that targeted plasmid DNA with high efficiency suggests that the problem is more likely related to inefficient packaging of plasmid DNA into virions in the sensor cells. Due to the extremely well characterized understanding of the lambda packaging sequences and associated mechanisms, it should be possible to optimize this approach to achieve a higher packaging efficiency of plasmid.

While this experiment (upon validation by subsequent replicates) could provide proof of principle for this approach, it is clear that optimization would be required before this approach would be viable. Additionally, the requirement that an entirely new plasmid be utilized in each

different sensor cell type with no sequence homology to any of the other plasmids used in any of the other sensor cell types is a fundamental limitation of this approach that would be nice to avoid if possible.

3.3.3.1.2 Chi site-directed specificity in the CRISPR adaptation response

The second approach we conceptualized to introduce unique information into signals generated by different sensor strains was to introduce specificity to the CRISPR adaptation machinery to target it to a specific site in the lambda genome. If this were possible, we could introduce a short segment of unique information (e.g. a DNA barcode) to that specific site and simply modify that information in the phage generated by each different sensor strain. The advantages of this approach include the fact that it would be highly scalable and require relatively small resource investment for each new sensor strain.

To accomplish this, we employed the use of χ (chi; crossover hotspot instigator) sites (Smith, 2012; Wigley, 2013). A chi site is a short (8 bp) DNA recognition sequence that regulates activity of the recBCD DNA repair complex in *E. coli* (Smith, 2012). The recBCD complex is a DNA repair complex that binds to double stranded DNA breaks in *E. coli* (Smith, 2012). The recBCD complex binds to double stranded DNA breaks and proceeds processively up the DNA strand until it encounters a chi site (Smith, 2012; Taylor & Smith, 1999; Wigley, 2013). Chi sites for recBCD function similarly to terminator sequences for polymerases; upon encountering the chi sequence, the recBCD complex stalls and eventually dissolves and dissociates from the DNA strand (Smith, 2012; Taylor & Smith, 1999; Wigley, 2013). As a result of interactions between the recBCD complex and the DNA at the chi site before dissociation, the site becomes primed for a DNA recombination (i.e. crossover) event at the chi site (giving

rise to its name, crossover hotspot instigator) (Smith, 2012; Wigley, 2013). The functions of chi sites are conserved across Gram-negative and positive species, although the exact recognition sequence differs between species (and in Gram-positive species the DNA repair complex is the addAB complex instead of the recBCD complex) (Wigley, 2013).

The currently accepted model for new spacer acquisition in CRISPR systems is that preprocessing of DNA by the recBCD complex is required to generate substrate that can subsequently be targeted by the CRISPR adaptation machinery (Figure 3.1 top right) (J. McGinn & L. A. Marraffini, 2019). A study in the Gram-positive species *Staphylococcus aureus* using the ϕ12 phage showed that during the CRISPR adaptation response, new spacers were preferentially acquired from chi site adjacent regions of the phage genome, and that chi sites acted to limit the region of genome sampled by the adaptation machinery for spacer acquisition (Modell, Jiang, & Marraffini, 2017). We therefore proposed that we might be able to insert *E. coli* chi site(s) into the lambda phage genome to impose specificity for the CRISPR adaptation machinery to chi site-adjacent loci. We used lambda red recombineering (see Materials and Methods section 3.2.15) and a kanamycin selection cassette to introduce 4 tandem chi sites into the lambda phage genome just upstream of the right end cos site as that end is always injected first into infected cells.

Integration of the chi site containing construct appears to have been successful as in two separate experiments, significantly more kanamycin resistant colonies were generated from our chi-site containing construct than a control construct without chi sites. Due to the natural function of chi sites as crossover hotspot instigators, this (along with kanamycin resistance) indicates successful introduction of our construct into the phage genome. A single

colony was selected, cultured under antibiotic selection, and induced to produce lambda phage virions. These virions were competent to successfully infect naïve *E. coli* K12-MG1655 indicator cells. These plaques had small colonies of stably lysogenic cells at their center. The lysogens from an entire plate were harvested and grown in liquid culture with antibiotic selection and then plated onto LB and kanamycin plates, resulting in a lawn of cells. The generation of kanamycin resistant colonies from lysogens resulting from a plaque infection experiment using virions from recombineered colonies strongly suggests that we were successfully able to introduce a genetic construct containing chi sites and a kanamycin selection marker into the lambda phage genome. Furthermore, these results strongly suggest that this recombinant phage genome is capable of induction from the lysogenic state, production of recombinant phage genomes which are packaged into virions followed by cell lysis, and that the resulting virions containing recombinant phage genome are competent for infection of naïve cells, resulting in new stable lysogens harboring the recombinant lambda phage genome. We look forward to performing a co-culture experiment with this recombinant strain and investigating the distribution of new spacer sequences along the lambda phage genome. It is our expectation that the sequences of new spacers will preferentially map back to lambda genome sequences that are adjacent to the newly incorporated chi sites.

3.3.3.2 Recording information correlated to sensing events as the evolution fluorescent activity in the memory population

3.3.3.2.1 Construction of recombinant fluorescent lambda phage

There are specific advantages to the sequencing-based read-out we have discussed so far. Specifically, because new spacers are always added to the same side of the CRISPR locus,

chronological information regarding the order of spacer addition is inherently encoded into the CRISPR array. However, it is not always optimal for resource/time intensive readouts like deep sequencing to be required for decoding information from biosensor devices. In many situations there are neither the resources nor personnel to perform these types of analytic pipelines. Another important dimension can be time; there are situations when it is not optimal or even viable to have to wait for days/weeks to decode the information from a memory component. Finally, it is interesting to consider the compatibility of our system with a more functional genetic payload (i.e. a cassette for expression of a gene of interest) as opposed to simply information transfer. We therefore endeavored to explore alternative readout modes that would allow for low resource decoding, real time feedback of signaling events, and or transfer of functional gene cassettes.

The most obvious and commonly used real time reporters in biological systems are fluorescent proteins. Fluorescent protein expression can often be used as a dynamic readout parameter to measure any biological activity or process that can be linked to protein expression, through any variety of transcription, translation, post transcription, translation, post-translation, protein-protein interactions, etc. We therefore sought a way to couple our distributed biosensor architecture with a real-time fluorescent protein readout, which could also conveniently be used to demonstrate compatibility of our system with transfer of functional gene cassettes.

The approach we designed is to clone a fluorescent protein expression into the lambda phage genome of our sensor cells. When inducer is added to the sensor cells (using our same previously discussed biosensor component), the recombinant lambda phage genome encoding

the fluorescent reporter cassette will be copied and packaged into virions that can subsequently infect memory cells, delivering the fluorescence cassette to the memory cells for expression of the fluorescent marker. Evolution of fluorescent signal can be used as a real time readout of a signaling event.

To implement this design, we first introduced a targeting construct with a fluorescent protein expression cassette under a tetR promoter and a kanamycin selection cassette into the lambda prophage in our sensor cell lysogens. To initially determine whether our fluorescence marker would be expressed strongly enough to be detected from its integration site in the lambda genome, we used lambda red recombineering to insert our fluorescence cassette into an *E. coli* K124 lysogen (a commonly used lambda phage chassis strain). A single colony from this recombination experiment was cultured under antibiotic selection overnight and was assayed for fluorescence compared to a control non-fluorescent strain. As seen in Figure 3.8, the resulting recombinant strain exhibits a strong fluorescent signal at the canonical green fluorescent protein wavelengths whereas the control strain does not.

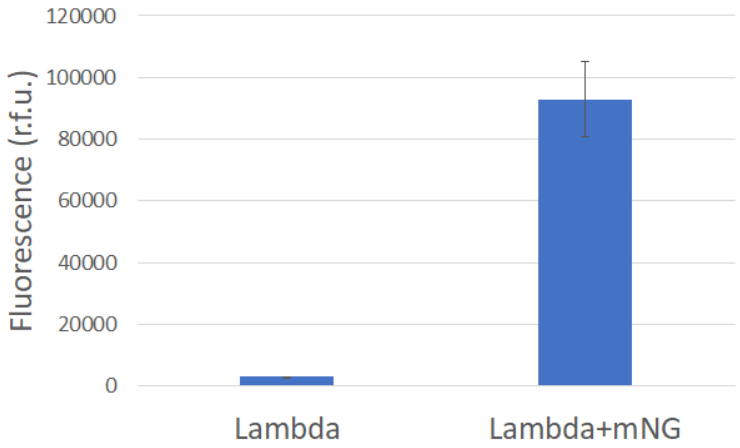


Figure 3.8 Bar graph depicting population level fluorescence detected from sensor cells encoding a recombinant lambda phage genome expressing mNeonGreen, a GFP variant. Population level fluorescence (right) is compared to fluorescence detected in a control strain (left) (error bar from overnight cultures of three separate colonies; fluorescence detected via Biotek platereader).

3.3.3.2.2 Future directions

Having established that our fluorescent protein is expressed strongly enough to be detected by population level fluorescence in a plate reader, the next important step is to demonstrate that this recombinant phage genome is competent for induction by our biosensor component, packaging into virions, and that those virions are capable of infecting host cells. This can be confirmed for the strain developed above using techniques already discussed (similar to experiments described in section 3.3.3.1.2).

Next we will introduce our recombinant fluorescent lambda genome into a sensor cell chassis strain that constitutively expresses the tet repressor from the genome, *E. coli K12 MG1655-Z1* (Sekar, Gentile, Bostick, & Tyo, 2016). Consequently, in sensor cells the fluorescence cassette integrated into the lambda genome (controlled by the tetR promoter) will be repressed by the constitutively expressed tet repressor. When these sensor cells are induced by signal, the recombinant lambda will be packaged into virions and subsequently infect the memory cells. Once inside the memory cells, which do not express the tet repressor, the fluorescence cassette will be expressed resulting in the evolution of a fluorescent signal from the community and allowing for the use of fluorescence evolution to be used as a readout of signal introduction. It is possible we will have to express the lambda phage repressor cI from plasmid in the memory cells to prevent the majority/all of the cells entering the lytic cycle after infection. Although with this system the ability to record order of exposure of signals to the community is lost without constant monitoring, what is gained is the ability to pair the system with a real time readout that is accessible in low resource setting (especially if a colorimetric

readout is used as opposed to a fluorescent one). This system also demonstrates compatibility of our system architecture with a functional gene cassette payload.

3.4 Discussion

The ability to extract information from internal and external environments, encode the information into biologically compatible molecules, transmit and share the information with adjacent cells, and eventually store some types of information stably over generations is vital to the success and survival of microbial communities. Each of these capabilities is also desirable within certain biotechnological contexts. Biosensors have been used for a wide variety of diagnostic and therapeutic applications (Song, Xu, & Fan, 2006; Vigneshvar, Sudhakumari, Senthilkumaran, & Prakash, 2016). Similarly, biomemory systems have been applied to diagnostic and therapeutic applications (Kotula et al., 2014; Sheth & Wang, 2018). Systems to integrate information from biosensors and share that information with an output or memory component are currently underexplored but are beginning to receive attention (Silva & Boedicker, 2019). Here we have developed a novel synthetic biosensing and biomemory system based on bacteriophage-mediated information transfer.

Our system integrates the three components (biosensing, information transfer, biomemory) within a single architecture in a design that is theoretically capable of sensing and responding to a large number of unique environmental signals and recording their order of exposure, which is not true of any currently or previously proposed biosensing/biomemory system. Immediate applications of such a system include contexts such as a persistent, clandestine sensor in remote environments. Development of a complimentary microfluidic

device could potentially enable such implementation. Additionally, however, with its distributed sensing array combined with the centralized memory component, our system begins to address the challenge of integrating multiple streams of information into a single analog or threshold output. This capability, which is characteristic of other biological systems such as neurons, could be a significant enabling technology within the context of *in situ* therapy production and delivery.

Currently drug delivery is one of the most complex challenges facing many potential promising therapies (Council, 2014; Homayun, Lin, & Choi, 2019; Hooven, 2017). To deliver biologically active forms of therapeutic compounds to their sites of interest in therapeutically relevant concentrations is extremely challenging (Council, 2014; Homayun et al., 2019; Hooven, 2017). In pill form, compounds must make it through the incredibly harsh digestive system until they can be absorbed into the blood stream in the gut and subsequently be transported to their sites of activity without being degraded or converted into inactive forms (Homayun et al., 2019). Therapies injected into the blood stream face similar challenges in addition to getting through the blood-brain barrier for neural therapies, and with average lifespan increases neurodegenerative disorders are one of the fastest growing illnesses in the world (Hooven, 2017).

In situ production of therapeutic compounds by cells in various human microbiome communities is one way to address some of these challenges. Compounds produced by bacteria in the gut, for example, could be immediately absorbed into the blood stream for transport or could be used to target illnesses such as inflammation and irritable bowel syndrome in the gut. Recent findings in the gut-brain axis have shown that pathogenically

relevant compounds produced in the gut are delivered to the brain via the vagal nerve, opening the possibility of delivery of neural therapeutics from gut microbiome produced compounds (Kim et al., 2019). Bacteria have also been engineered to target cancer tumors in certain contexts (S. Zhou, Gravekamp, Bermudes, & Liu, 2018).

In most of these scenarios, many of the components are already developed or relatively close to development, with one glaring exception. Each of these drug delivery systems would require a robust, stable, control system capable of integrating a variety of biological signals to diagnose a condition into a single response element to produce a therapy. Currently such systems do not exist. Our distributed sensing network coupled with a centralized memory components provides a framework for a way to begin thinking about what systems that can integrate multiple streams of information from different sensor sensor into a single output might look like.

Chapter 4: Conclusions and Future Directions

4.1 Review of motivations

Synthetic biology is a field which combines molecular biology with technology and engineering to facilitate and accelerate the design of biological systems for medical, industrial, and environmental applications (Shapira, Kwon, & Youtie, 2017). Although generally agreed to have emerged around the turn of the millennium, the roots of synthetic biology reach back to advances in the 1970s (and further) when the tools for what we think of as modern genetic engineering were first being discovered and explored. Following several grand medical and monetary biotechnology success stories such as the cloning of the human insulin gene into a bacterial plasmid for the microbial production of human insulin, investment of time, money, and resources into the field have been steadily increasing (CBInsights, 2019; Chen, Liu, & Chen, 2009; Medicine, 1995; Shapira et al., 2017). As the field has progressed, the biological complexity and scale of systems of interest has also increased significantly, from the expression of single gene targets in model organisms to the expression of biosynthetic pathways containing tens of genes, at times requiring macromolecular assembly and organization (Dutta et al., 2014; Paddon & Keasling, 2014; Sachdeva, Garg, Godding, Way, & Silver, 2014). Many orthogonal tools and technologies have been developed to accommodate the requirements imposed by this increase in complexity.

In the first decade of the new millennium, a novel approach was proposed to address many of these challenges simultaneously. Instead of choosing a single model organism to express all components and processes of target systems, it was suggested that systems could be distributed amongst several different species and/or cell types within a microbial community. There are a variety of potential advantages and challenges to this type of approach which have been discussed at length in chapters 1-3.

4.2 Specific contributions to the field of synthetic biology

Throughout the course of my thesis work, my goal has been to develop tools and strategies to address some of the challenges of engineering microbial communities that we might better be able to access their advantages. Specifically, chapter 2 focuses on developing and characterizing a tool that can be used to enable species coexistence and to program relative species abundance in a synthetic microbial community. Tools to accomplish this task are important because programming this parameter of synthetic microbial communities can be a useful approach to optimize community performance for biotechnological applications.

Chapter 3 focuses on the development and characterization of a novel synthetic microbial community with a distributed sensing and centralized memory architecture that relies on bacteriophage-mediated information transfer to sense and record the environmental stimuli to which the community was exposed. This project introduces not just a novel system but a novel theoretical approach within the field, combining a distributed sensing network with a centralized memory component. This approach is theoretically significantly more amenable to

a variety of potential therapeutic applications of synthetic microbial communities in the human microbiome.

4.2.1 Compare and contrast to existing technologies

4.2.1.1 Temperature regulation as a tool to program synthetic microbial community composition

The ability to control synthetic microbial community composition is desirable within a biotechnological context. Until recently, this issue has gone almost completely unaddressed within the synthetic biology field. However, within the past three years, several studies have been published that demonstrate a variety of approaches to address this issue. A pair of studies published by Scott et al. in the Hasty lab at the University of California, San Diego in 2017 and Stephens et al. in the Bentley lab at the University of Maryland in 2019 used quorum sensing to control expression of genes that controlled cell death and cell growth, respectively, to control community composition in synthetic microbial communities (Scott et al., 2017; Stephens et al., 2019).

Briefly, the paper from Scott et al. placed a circuit regulating expression of a lysis gene under regulation of a quorum sensing circuit. When the population density of one strain reaches a designated level, expression of the lysis circuit is induced, resulting in cell death. This self-limiting control strategy is used to affect population control and regulate community composition. This concept is a variation on a theme published by You et al. in 2004 for population control by quorum sensing mediated cell killing (You, Cox, Weiss, & Arnold, 2004).

One advantage to the community composition control system published by Scott et al. is that because there are many well characterized, orthogonal quorum sensing systems, their approach can theoretically scale well to communities with many more than just two species. However, the main drawback to the system developed by Scott et al. (and by You et al.) is that cell death imposes a very strong evolutionary selection pressure to escape whatever genetic circuit is controlling the cell death. Consequently, both the Scott et al. and You et al. systems are plagued by very short culture lifetimes, because relatively quickly mutations emerge to escape expression of gene resulting in cell death.

In comparison to the system developed by Scott et al., our system has strengths and weaknesses. The major weakness of our system is that the control mechanism does not likely scale well. It should be possible to develop a temperature control regime for up to three species, but developing a control regime for any additional species would be a daunting task. However, the major strength of our approach is that because response the temperature is such a deeply ingrained, hard-wired response in bacteria, evolution and adaptation to minimize responses to temperature is very slow. We were able to demonstrate that our systems were functional for up to 400 hours (constant temperature) and 96 hours (cycling temperature). And even in the event of adaptation to temperature changes which affects the rate at which bacteria adapt to temperature changes, the response to temperature overall can be maintained for hundreds or thousands of generations (Bennett, Lenski, & Mittler, 1992).

In 2019 Stephens et al. published a system that is a relatively fresh approach to using quorum sensing to control community composition. Instead of using quorum sensing to control cell death, they used quorum sensing molecules to regulate expression of a gene that increases

cell growth. Because the phenotypic behavior controlled by the synthetic circuit in this case is increased growth, the circuit developed by Stephens et al. is significantly less likely to be quickly escaped by mutation. However, overexpression of the growth factor in their system represents a significant diversion of resources away from any biosynthetic production pathway of interest, likely resulting in a consequent negative impact on yield in a production strain. Additionally, despite the fact that their system regulates cell growth instead of cell death to control community composition, it is not clear whether their system would last as long as our temperature-based system before mutations emerge in their circuitry. Their system has the same strength as the system from Scott et al., however, which is that due to the number of orthogonal quorum sensing systems available, their system potentially scales well to larger numbers of species.

Therefore, the weaknesses of our system compared to that of Stephens et al. are the same as with Scott et al., that scalability of our system is poor compared to quorum sensing based systems. The strength of our system compared to Stephens et al. is that we do not rely on overexpression of any cellular component that would significantly divert resources away from the biosynthetic pathways of interest, and the lifetime of our control system is likely longer.

4.2.1.2 Distributed sensing with centralized memory utilizing bacteriophage-mediated information transfer in synthetic microbial communities

A variety of biosensing and biomemory systems have been previously developed and demonstrated in the literature. Overwhelmingly the approach has been to use a signal cell to express all components of the system, including any and all sensor and memory related

components (Farzadfard & Lu, 2014; Kotula et al., 2014). The most recent, and best, iteration of this approach was published by Sheth et al. in 2017 (Sheth et al., 2017). In their ingenious system, Sheth et al. use the CRISPR adaptation machinery as a memory component to record DNA barcodes that are generated in response to biosensor activation. However, because the Sheth et al. system uses a single cell architecture, they face a choice to record the order of exposure of signals for a large number of biosensors. Either they can insert as many biosensors as possible into a single cell line (centralized sensing with centralized memory), or they can use a number of different cell lines, each with a single biosensor and memory component, but include a time keeping mechanism so that exposure profiles of the different cell lines can be compared to one another, producing an exposure profile for the community (distributed sensing with distributed memory). Sheth et al. opted for the latter, and demonstrated their approach works (mostly) with three biosensor cell types, but due to the nature of their system design and the problems encountered with only three sensor strains, it is unlikely their system can scale significantly beyond that number.

Our system represents a significant improvement in this regard. Due to our distributed sensing system combined with a centralized memory component, our system can theoretically incorporate a large number of sensor cell lines while preserving the ability to record the order of exposure to signals for the entire community. Additionally, our separation of the sensing and memory components enables another improvement over the system from Sheth et al. The time keeping device used in the Sheth et al. system relies on constitutive overexpression of the CRISPR adaptation complex Cas1-Cas2, leading to genome instability in the sensor/memory cells. Because our system design does not require this feature, we can theoretically dispense

with this constitutive overexpression by using a quorum sensing regulated system to only express the CRISPR adaptation machinery when necessary. This would theoretically increase the maximum culture lifetime of our biosensor/biomemory system significantly over Sheth et al.

The weakness of our system is that because we have separated the sensory and memory components into different populations, we are forced to develop a novel information transfer system. It has been difficult to optimize a variety of aspects of this design, first and foremost the ability to encode unique information into the signals produced by each sensor cell strain. However, upon surmounting this obstacle our system will be superior to previous systems in multiple ways. Additionally, the intellectual shift from thinking about these biosensor systems as single cell systems to distributed systems enables a new way of thinking about potential applications for biosensor/biomemory systems, which is exciting.

Both of these projects add to the toolbox available to researchers to manipulate and program coordinated behavior within synthetic microbial communities. The naturally emerging next question is, therefore, for which types of applications are the capabilities of synthetic microbial communities best suited? Which of the multitude of pressing challenges currently facing the human species are synthetic microbial communities best suited to address?

4.3 Future directions

4.3.1 Temperature regulation as a tool to program synthetic microbial community composition

4.3.1.1 Feedback control for temperature regulation

In Chapter 2 we developed temperature regulation as a tool to enable coexistence and program community composition within a synthetic microbial community. We developed two separate techniques to use temperature regulation to program community composition, which we refer to as constant and cycling temperature regimes. In the cycling temperature regime approach, one challenge we encountered was biological adaptation to repetitive exposure to stimuli. It has been documented that some bacteria are able to develop transient and/or sustained adaptations to cyclical exposure to certain stimuli (Forbes, Dobson, Humphreys, & McBain, 2014; Fridman, Goldberg, Ronin, Shoresh, & Balaban, 2014). We also observed variations in the lag times and overall response to temperature changes in some of our cycling temperature regime experiments (Figure 2.4, Figure 2.6). Even relatively small variations in these parameters of the adaptation response to temperature changes make it difficult to calculate *a priori* a sequence of time intervals at specific temperatures to achieve a certain community composition. However, the phenotypic response to temperature is deeply hardwired into bacterial organisms, persisting for hundreds and even thousands of generations (Bennett et al., 1992).

One approach that would theoretically allow us to circumvent these variations in the time course of the response of bacteria to temperature change while still enabling the use of temperature regulation to program community composition is incorporation of a feedback

control unit. The basic design for such a control system is relatively straightforward. One design for such a system would require two components, a community composition monitoring component, and a temperature control component. For monitoring community composition, in our system it would be possible to simply use fluorescence, either at the population level or by real time flow cytometry monitoring (Broger, Odermatt, Huber, & Sonnleitner, 2011). A software interface with this monitoring system would allow the user to set a desired target community composition. If and when the measured community composition were to deviate from the target composition, this information would be forwarded to the temperature control component, which would adjust the temperature in such a way so as to return the community composition back to the target.

While simple in design, there would almost certainly be non-trivial optimization required for this type of a feedback control system. One crucial parameter to optimize would be the amount by which the temperature is changed in response to deviations in community composition, and the frequency with which these temperature steps are applied in response to deviations in composition. However, in theory this feedback system would enable a temperature regulation approach to maintain a wide variety of target community compositions in spite of variations in the time course with which species response to temperature changes.

An additional benefit to this approach would be that for many industrial sized culture volumes (which are sometimes in the tens of thousands of liters), the amount of energy and therefore money required to change the temperature of the culture media is significant. Implementing a real-time feedback system as described above would allow the system to

minimize the temperature changes required to maintain the target composition, potentially resulting in significant environmental and monetary improvements.

4.3.2 Quorum sensing-controlled expression of CRISPR adaptation machinery

In chapter 3 we developed a novel biosensor/biomemory system that uses a distributed sensing array with a centralized memory component and bacteriophage-mediated information transfer. In the system we have developed, we co-opt the use of the CRISPR adaptation response based memory component developed by Sheth et al. (Sheth et al., 2017). One aspect of the component as developed by Sheth et al. is that the CRISPR adaptation machinery are constitutively expressed. This feature is required in their system as it is part of the time-keeping device they use that is necessitated by their distributed memory approach. However, as discussed previously, over expression of the Cas1-Cas2 CRISPR adaption complex at the levels required for this application causes the complex to lose its specificity for foreign DNA and attack the *E. coli* genome in addition to any foreign DNA fragments. Unsurprisingly, this leads to genome instability and significantly limits the stability of the system, resulting in unstable cultures that only last days/weeks.

However, because our system uses a centralized memory component enabled by our bacteriophage-mediated information transfer, we do not require the time keeping device that Sheth et al. relied on and can therefore avoid the constitutive expression of the CRISPR adaptation complex. The question becomes, then, is there a better expression scheme that would enable greater system stability?

One approach that would capitalize on our distributed sensing and centralized memory system architecture is to regulate expression of the CRISPR adaptation machinery via quorum sensing. The basic system design is depicted in Figure 4.1. As seen in Figure 4.1 A, the system

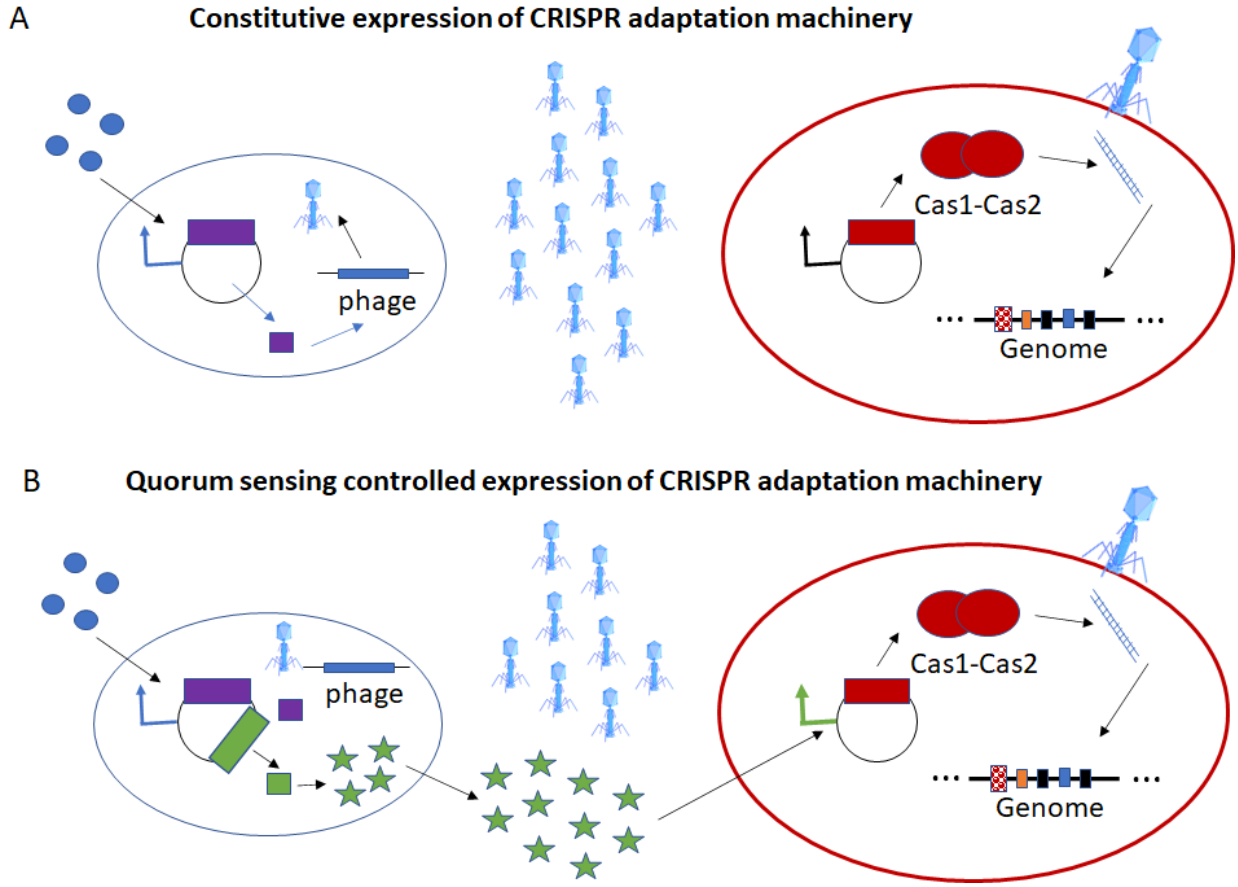


Figure 4.1 Cartoon depiction of the putative scheme to control expression of the Cas1-Cas2 complex with quorum sensing. A) the biosensor/biomemory system described in chapter 3. B) modifications required for quorum sensing controlled expression of CRISPR adaptation machinery. LuxI, the AHL synthase is depicted by a green square. The quorum sensing signaling molecule AHL is depicted by green stars. The quorum sensing controlled promoter pLux is depicted by a green promoter symbol. Production of LuxI (and therefore AHL and eventually the Cas1-Cas2 complex) is initiated by presentation of biosensor inducer (represented by small blue extracellular circles).

from Sheth et al. utilizes a constitutive promoter for the Cas1-Cas2 complex (depicted by the black promoter symbol). Under the proposed quorum sensing regulated system, expression of Cas1-Cas2 would be placed under the quorum sensing promoter pLux (green promoter symbol in Figure 4.1 B). Transcription initiation from the pLux promoter is controlled by the regulatory protein LuxR, an inducible transcription factor. In the absence of its inducer, AHL (depicted by

green stars in Figure 4.1 B), LuxR will remain inactive, and expression of genes downstream of the pLux promoter is repressed. In the presence of AHL, LuxR binds AHL and assumes a conformation which allows it to interact with the pLux promoter, initiating transcription at the promoter.

For quorum sensing controlled expression of the CRISPR adaptation machinery, all that is required would be for the AHL synthase LuxI to be cloned into the operon downstream of recA730 in our sensor cell strain (green rectangle in Figure 4.1 B). With this circuit design, in response to sensor strain inducer, not only would recA730 be expressed to induce the lytic cycle of lambda phage but the LuxI synthase would also be expressed, resulting in production of AHL. The diffusible AHL would then interact with LuxR and activate expression of the CRISPR adaptation machinery in memory cells, which would then target the injected phage genomes from the sensor strain.

The benefit of this expression system is that it would act to extend the culture lifetime of the system indefinitely. Constitutive expression of the CRISPR adaptation machinery results in genome instability and culture collapse, but intermittent, pulse expression of the adaptation machinery only when a signal is encountered by the community would avoid this genome instability. This quorum sensing regulated control of the CRISPR adaptation machinery would therefore capitalize on a unique feature of our system to address a significant limitation of a previous approach.

4.3.3 *In situ* temporally and spatially appropriate microbial production of therapeutics in the human microbiome

4.3.3.1 Challenges in therapeutic drug delivery

Synthetic microbial communities have been and should continue to be used to address challenges across a variety of medical, industrial, and environmental applications. One current significant challenge in the medical care field is the delivery of clinically relevant concentrations of active therapeutic compounds to the site of action (Council, 2014; Homayun et al., 2019; Hooven, 2017). One of the most common routes of administration, oral delivery, must contend with problems including fabrication protocols that avoid use of organic solvents, shear stresses, and local temperature increase to avoid drug inactivation, in addition to which delivery material, design, size, and polydispersity must be accurately controlled due to their influence on treatment efficacy. Additionally, designing drugs that can survive the route of delivery and still effectively cross various biological barriers such as the lumen, mucus, tissue of the GI tract, and the blood brain barrier can be difficult. A relatively new approach to these problems is to genetically engineer living bacteria in the human microbiome and cancer tumor microenvironments as diagnostic and therapeutic agents (Charbonneau, Isabella, Li, & Kurtz, 2020; Lim & Song, 2019).

There are a number of technological and ethical concerns with respect to these approaches, but given the potential advantages to *in situ* production and delivery of therapeutics and recent advances in related technologies, it is currently appropriate to begin thinking about the practical and ethical hurdles inherent to such approaches (Charbonneau et al., 2020). In fact, some bacteria have already been successfully engineered for such

applications and are currently in phase I, II, and III clinical trials (Lim & Song, 2019; Riglar & Silver, 2018). Currently a number of therapeutic agents are already developed and waiting to be paired with appropriate bacterial production chassis and regulatory control systems for *in situ* therapy production and delivery.

4.3.3.2 *In situ* microbial production of therapeutics

One system architecture that could potentially enable *in situ* production of therapeutic would be a system of therapeutic production strains that are governed by a distributed system of biosensor strains. The production strains would exist as therapeutic nodes and would encode biosynthetic pathways responsible for production of therapy that addresses a specific pathogenic state. Activation of these biosynthetic pathways and production of therapy would be controlled by signals sent from any number of sensor strains which would be activated by biomarkers and/or environmental conditions associated with the relevant pathogenic state. For implementation of such a system, several looming questions must be addressed. Among the most important of these are 1) what might an optimal information transfer mechanism from the sensor cells to the therapeutic nodes look like and 2) what might some of these therapies look like?

4.3.3.2.1 Membrane-bound vesicle-mediated information transfer

In the system we developed in Chapter 3, we pioneered a bacteriophage-mediated information transfer system. When thinking about system requirements for an information transfer mechanism within our *in situ* microbially produced therapy context, there are some advantages and disadvantages to a bacteriophage-mediated information transfer system. One of the advantages is that bacteriophages are generally highly specific in terms of the target cells

they are capable of interacting with. The tail fibers of a specific species of bacteriophage are generally only capable of interacting with a single or a very small range of receptor types (Yehl et al., 2019). Within this context that is an advantage because depending on the payload of these phage, it might be important that the signal generated by sensor strains be delivered only to the therapeutic node cells. However, in this context that specificity might also be a disadvantage, for two reasons.

First, a range of *E. coli* strains are known to be commensal residents of the human gut microbiome, and it might be important to limit delivery of signal from sensor strains to these cells, or at least limit the signal dilution that might occur. Secondly, it is not currently clear what model organisms will be optimal for the therapeutic production strains. For many of the reasons discussed throughout this thesis, it is entirely possible that a variety of species might be desirable as production strain chasses. One final disadvantage of bacteriophage-mediated information transfer in this context is that generally only nucleic acid encoded information is competent for packaging into bacteriophages. While this is not necessarily always a disadvantage, a more optimal vector for information transfer in this context would be able to theoretically package a wider variety of compounds. Additionally, in this context it might be desirable to integrate information from multiple sensor cell strains into a single output in a therapeutic node (as many pathogenic states can only be correctly diagnosed by the simultaneous presence of several symptoms), and while not necessarily impossible, it is not immediately apparent how that might be accomplished with nucleic acid based information transfer. So, a more optimal information transfer vector in this context would be able to package a wide variety of compounds and would be able to deliver those compounds with

specificity to a wide variety of target species. On both of these counts, membrane bound vesicles offer opportunities that bacteriophage-mediated information transfer cannot.

Membrane bound vesicles bleb off from almost all known cell types and are one of the most conserved characteristics observed in cellular life (Deatherage & Cookson, 2012). While the molecular packaging mechanisms have yet to be elucidated, a wide variety of molecules ranging including periplasmic and cytosolic proteins, DNA, RNA, virulence factors, and quorum sensing molecules have been isolated in membrane bound vesicles (Deatherage & Cookson, 2012; Toyofuku, Nomura, & Eberl, 2019). Additionally, not only interspecific but also inter-kingdom vesicle mediated transfer of compounds has been observed (Cai et al., 2018). Finally, the combination of membrane bound vesicle mediated information transfer with quorum sensing controlled regulation of gene expression offer one path forward to integrate multiple information streams into a single output. A recently published study by Silva et al. demonstrate that a combination of different quorum sensing molecules and quorum sensing receptors can be used to create gene expression systems that are dependent on the community level signaling state as opposed to the signaling state of a single signaling system, which provides a mechanism to integrate multiple information streams into a single output (Silva & Boedicker, 2019). Vesicle mediated quorum sensing molecule transfer could be one way to implement this type of system within a human microbiome therapy production system.

While the characterization and optimization of vesicle loading mechanisms as well as the targeting mechanisms used to direct vesicles to specific cells or cell types that will be required to develop this modality as a viable information transfer mechanism in the context of

a human microbiome therapeutic community, in the long term membrane bound vesicle offer unique opportunities as an information transfer mechanism.

4.3.3.2.2 Therapeutic nodes

The next important question in regard to how *in situ* therapy production can be implemented is what the specific therapeutic nodes might look like. One excellent example of a strain that could be deployed as a therapeutic node was recently published by Hsu et al. and involves bacteriophage-mediated delivery of a repressor protein that prevents expression of a virulence factor in a pathogenic bacterium . While we have already discussed the potential role of bacteriophage as information carriers in our putative system (and more specifically their shortcomings with compared to the potential of membrane bound vesicles), it is now appropriate to discuss their role in therapy delivery.

As discussed previously, one of the key characteristics of bacteriophages are their target cell specificity. This characteristic has already made "phage therapy" an extremely active area of research. Under the phage therapy paradigm, pathogen specific bacteriophage are used to eliminate pathogens instead of antibiotic treatment in cases of life-threatening bacterial infection. Phage therapy offers distinct advantages over antibiotic therapy in that it can be effective even in the face of the evolution of antibiotic resistance and also in that it can target pathogens for killing much more specifically than antibiotics, which often target a wider range of species for killing, sometimes including important human microbiome symbionts.

However, even phage-mediated killing can result in the relatively quick evolution of bacterial resistance due to the fact that such a strong selection pressure is being applied

(Oechslin, 2018). Consequently, an emerging field of phage therapy involves phage mediated delivery of a genetic payload that targets the pathogenic virulence factor without killing the pathogen. This approach could theoretically greatly reduce the evolution of bacteriophage resistance as it removes one of the strongest pressures selecting for resistance (pathogen cell death). The previously mentioned therapy developed by Hsu et al. employs just such a virulence factor targeting therapy (Bryan B. Hsu, Way, & Silver, 2020). But their approach is not at all modular, and is only capable of targeting a very specific set of virulence factors.

A recent significant discovery has paved the way for an almost entirely modular system to target virulence factors in pathogens that is theoretically entirely deliverable by bacteriophage. A pair of papers published within a few weeks of each other in 2019 described a CRISPR guided transposon system (CAST system) that allows a genetic payload to be integrated into a genomic site of choice targeted by a CRISPR guide RNA. Because the entire system consists of less than 10 kb of genetic material, it is theoretically plausible for the system to be encoded and delivered to a pathogen via bacteriophage. The CAST can be programmed to target any specific virulence factor and insert a genetic payload that knocks out expression of the virulence factor. It might even be possible to use a genetic payload that would undergo slight positive selection to ensure reversion to wild-type does not occur.

However, there is still one significant problem with this therapy approach. As discussed previously, most bacteriophage are very specific. So, while the CAST system is easily re-targetable to different virulence factors, for each different pathogen a new bacteriophage chassis must be found. Often these bacteriophages are necessarily hosted in the pathogen strain of interest, and genetic tools to modify these non-model strains can be difficult,

expensive, and time-consuming to develop. For this reason, it would be optimal to create a modular phage chassis that is capable of packaging a genetic payload of interest (that does not necessarily include the phage genome but could include a CAST system) and that includes easily modifiable, modular tail fiber proteins to enable targeting of the phage chassis to different receptor types. As discussed previously, phage specificity is governed almost entirely by the phage tail fiber proteins, which largely determine the receptor(s) with which phage can interact.

Unsurprisingly, there are already attempts underway to use directed evolution to modify these phage tail fiber proteins to modify phage target cell specificity (Yehl et al., 2019). While directed evolution of existing biological systems is often an extremely powerful tool, it can also be limited by the recent evolutionary trajectory of said systems. Additionally, in many cases, billions of years of evolution have resulted in systems that favor efficiency over modularity, as evidenced by bacteriophage (Butterfield et al., 2017). In such cases, synthetic systems rationally designed from first principles can potentially provide a modular chassis that can then be subsequently evolved to generated desired properties for modular delivery of genetic payloads to a variety of cell types (Butterfield et al., 2017). Butterfield et al. used just such an approach to develop a synthetic, computationally designed nucleocapsid that packages its own full length mRNA genome (Butterfield et al., 2017). This type of an approach could theoretically be applied to generate a CAST system delivery vector that would be extremely modular and could be targeted to a wide variety of cell types.

4.3.4 Conclusions

In this thesis I have discussed two novel systems that address specific current challenges regarding the development of synthetic microbial communities for industrial, environmental, and medical applications. In one system we demonstrate that temperature regulation can be used to enable coexistence and program community composition in a synthetic microbial community, and in the second system we explore a novel biosensor/biomemory system architecture that uses a distributed sensing component coupled with a centralized memory component enabled by bacteriophage-mediated information transfer. These systems address current and emerging challenges relating to coexistence, community composition, and communication in synthetic microbial communities. Additionally, the distributed architecture and information transfer component of the second system lay the intellectual groundwork for potential therapeutic applications of the system. In total we have presented data and analysis that represents a significant contribution to the field of synthetic biology and that lays the foundation for future contributions.

Appendices

Appendix A

The Relationship Between Temperature and Relative Fitness Differences and Niche Differences in a Synthetic Microbial Community

A.1 Introduction

A.1.1 Brief history of coexistence theory

In chapter 2 we demonstrated that temperature regulation can be used to enable coexistence between species in a synthetic bi-culture and it can also be used to modulate species relative abundance in that community in a programmable way. We showed that the co-culture growth rates change as a function of temperature and that the relative abundance of species in consecutive batch cultures is correlated to these growth rate changes. However, as mentioned previously and as illustrated in Figure 2.3 B, the species relative abundance at equilibrium in the bi-culture responds to temperature in a highly nonlinear manner. This observation suggests that more factors influence relative species abundance in the bi-culture than just the changes in growth rate in response to changes in temperature, which is an idea that is compatible with over a century of coexistence theory by natural and mathematical ecologists. As the mechanisms influencing coexistence of species in communities have been relatively unexplored by synthetic biologists, for a deeper understanding of how and why the relative abundance of species in our bi-culture responded to temperature changes as it did we turned to the wealth of coexistence research produced by the field of ecology since the early 20th century.

Modern coexistence theory arose in large part from the competitive exclusion principle, which is commonly attributed to Gause but is a consequence of the dynamic model of competition proposed by Volterra in 1931 (Gause, 1934; Godwin, Chang, & Cardinale, 2020; Volterra, 1928). The competitive exclusion principle can be briefly stated as "complete competitors cannot coexist" (Hardin, 1960) and contends that no two species with identical niche requirements can coexist indefinitely (Gause, 1934). Within this context, niches are commonly defined as responses of an organism or population to the distribution of resources and competitors, and when two species differentiate their niches, they often compete less strongly with each other (i.e. the effect of intra-specific competition rises relative to the effect of inter-specific competition) (Beals, Gross, & Harrell, 1999).

The impact of the competitive exclusion principle was such that, for a time, much of the work produced by the field focused on characterizing the impact of niche differences on community population dynamics. Concurrent to this focus on the role of niche differences in coexistence theory, however, was theoretical work which suggested that niche differentiation by itself was insufficient to fully explain coexistence in several cases (Godwin et al., 2020; Macarthur & Levins, 1967; Roughgarden, 1976; Slatkin, 1980; Vandermeer, 1975). The field would have to wait until the turn of the millennium for this inconsistency to be resolved.

In 2000, mathematical ecologist Peter Chesson published a synthesis of modern coexistence theory that provided a mathematically rigorous framework for the integration of many of the extant concepts into a single paradigm (Chesson, 2000). Within Chesson's framework all coexistence mechanisms can be separated into two broad classes, which he referred to as equalizing forces and stabilizing forces (Chesson, 2000). Equalizing forces are

defined by Chesson as those mechanisms that minimize average fitness differences between species, where average fitness differences are a function of species specific biological traits such as potential growth rates, carrying capacities, and resistance to consumers (Chesson, 2000; Godwin et al., 2020). Equalizing forces were initially referred to as average fitness differences by Chesson but have more recently been termed relative fitness differences (RFDs), and will be referred to as such throughout the remainder of this thesis (Godwin et al., 2020). Stabilizing forces were defined by Chesson as those mechanisms that tend to increase negative *intra*specific interactions relative to negative *inter*specific interactions (Chesson, 2000). Stabilizing forces can therefore be more easily conceptualized as niche differences (NDs) and have come to be commonly referred to as such and will be for the remainder of this thesis (Godwin et al., 2020). Chesson showed that under most models of competition, whether or not coexistence will occur is ultimately determined by the relative magnitudes of the applicable RFDs and NDs (Chesson, 2000; Godwin et al., 2020). Specifically, Chesson developed a coexistence criterion in the form of an inequality (denoted here according to notation from Carroll et al. 2011 and Narwani et al. 2013 (Carroll, Cardinale, & Nisbet, 2011; Narwani, Alexandrou, Oakley, Carroll, & Cardinale, 2013)

$$1 - ND < RFD < \frac{1}{1 - ND}$$

and proposed that coexistence can only occur when the magnitudes of the RFDs and NDs in the system are such that the inequality is satisfied (Chesson, 2000; Godwin et al., 2020). It is important to note here that Chesson's theory is not a new model of species interaction but

instead is a framework that can be generalized across a wide variety of competition models to predict coexistence, nor did Chesson initially propose an empirical method to quantify RFDs and NDs in practice (Godwin et al., 2020). However, there are at least two important aspects of the framework that set it apart from previous efforts. The first is that it elegantly reduces the dimensionality of a variety of models down to two terms (RFD and ND) which have clear ecological interpretations (Godwin et al., 2020). The second is that due to the nature of those two terms, knowledge of the specific competition mechanism(s) active within a given system is not required to quantify the RFDs and NDs. This brings us to the subject of RFD and ND quantification.

A.1.2. Quantification of RFDs and NDs

Quantification of RFDs and NDs is central to this entire exercise as the magnitudes of the RFDs and NDs become the final arbiters for coexistence prediction within Chesson's framework. And while as indicated above, knowledge of specific competition mechanisms within the system is not required to quantify RFDs and NDs, this is not to say quantification is necessarily trivial. Since the proposal of the framework, a number of approaches have been developed to quantify RFDs and NDs. Many of these methods have been derived from different models of species interaction and therefore require different assumptions and different study designs to obtain the relevant data (Godwin et al., 2020). Recently an in-depth analysis has been published which evaluates the different strategies and the extent to which each is appropriate for specific experimental systems (Godwin et al., 2020). Of the approaches discussed by Godwin et al., there are two which are generally most appropriate for liquid microbial cultures, the sensitivity method and the method based on the Lotka-Volterra model for continuous

reproduction. The sensitivity method calculates the RFDs and NDs by comparing the growth rate of each species in monoculture to the growth rate of that same species trying to invade a steady-state culture of the other species (Godwin et al., 2020).

The relative reduction in growth rate of a species in monoculture compared to the growth rate of that same species invading another at steady state is used to quantify the sensitivity of that species to interspecific competition, and it has been shown that NDs are proportional to the geometric mean of the sensitivities of each species whereas the RFDs are represented by variation around the mean (Carroll et al., 2011; Godwin et al., 2020; Narwani et al., 2013). Despite the enviable and elegant simplicity of this approach, certain systems are more amenable to quantification of RFDs and NDs via sensitivity analysis than others. Specifically, monoculture populations must be maintained at steady-state population for the invasion assays. While this is often achievable for relatively slower growing organisms such as algae, it is less straightforward for certain relatively faster growing species of bacteria, often requiring chemostat conditions which are not available in the Lin lab. Because the experimental approach required for the sensitivity analysis is not easily implementable in the Lin lab, we looked more closely at the other RFD and ND quantification method appropriate for microbial liquid cultures, the method based on the Lotka-Volterra model for continuous reproduction.

To quantify RFDs and NDs with the method based on the Lotka-Volterra model for continuous reproduction, high resolution growth curves of species in both monoculture and co-culture must be available. While generating growth curves of sufficiently high resolution is often painstakingly time consuming for some slower growing organisms such as algae, in the

case of some relatively faster growing species of bacteria it is relatively straight forward. Species can be inoculated into 200 uL volumes in a 96-well plate and cultured for ≤24 hours in a platereader which can automatically take optical density and/or fluorescence measurements at the desired interval (every 10 minutes, every 1 minute, etc.). The resulting growth curves are often of sufficiently high resolution and quality that the Lotka-Volterra continuous reproduction equations can be fit to the data to estimate the interspecific competition coefficients.

The RFD and ND quantification method based on the Lotka-Volterra model for continuous reproduction therefore best compliments the experimental approaches available with the model system we developed in Chapter 2, and so we decided to implement this workflow to interrogate our system that we might better understand the ways in temperature affected coexistence and species relative abundance within the framework of modern coexistence theory. Our expectation was that the insights gained from this approach would provide a deeper understanding of the coexistence mechanisms at work in our model system from Chapter 2 which may augment our ability to control that system and in addition contribute to the quickly adapting body of work on modern coexistence theory by interrogating the effect of temperature on RFDs and NDs and coexistence.

To accomplish this, we cultured both species from our synthetic microbial bi-culture in chapter 2 in monoculture and co-culture, then used the method based on the Lotka-Volterra model for continuous reproduction to estimate the intra- and inter-specific competition coefficients and from those the RFDs and NDs for our system across a range of different temperatures. We then looked for any correlation between the RFDs and NDs and temperature changes.

A.2. Materials and Methods

A.2.1. Generation of growth curve raw data

E. coli K12 substr. MG1655-YFP and *P. putida KT2440*-mCh were seeded from -80 ˚C cryostocks and grown in monoculture overnight to stationary phase in 14 mL Corning Falcon® test tubes (polypropylene test tube, round bottom, 17x100mm, 14mL, graduated, with clear snap cap, Sterile, 25 per Pack) in a 2 mL volume of M9 minimal media. Overnight cultures were diluted 1:100 into fresh M9 media and grown into exponential growth phase (OD600 ~0.4-0.6). The cell density of the two cultures of exponential phase cells were normalized to each other using OD600 and then diluted 1:100 into fresh M9 media in a Greiner Bio-one CELLSTAR™ 96 well μClear flat bottomed microplate with a 200 uL final volume. Strains were grown in monoculture and co-culture in triplicate (2 uL of the desired strain was added for monocultures, and 1 uL each of both strains was added for co-cultures to 198 uL M9 media). These nine wells were surrounded by wells filled with 200 uL sterile deionized H2O to inhibit evaporation from experimental wells. The lid was treated with a mixture of 20% ethanol + 0.5% Triton X-100 to avoid condensation formation on the lid (pour enough mixture to completely cover the bottom of the lid, let sit 5 minutes, pour off and let air dry). The lid was taped on using Fisherbrand™ labeling tape. The plate was incubated in a Biotek Synergy H1 platereader for 24 hours with plate reads every 10 minutes and continuous orbital shaking at 282 cpm. At each timepoint, reads at 600 nm wavelength, Excitation: 510 nm Emission: 540 nm, and Excitation: 585 nm Emission: 620 nm were taken of each well (these wavelengths were empirically determined to maximize the specific signal and minimize crosstalk between the two channels for our constructs, media, and platereader; data not shown).

A.2.2. Raw data processing

A.2.2.1 Stationary phase adjustment

One of the drawbacks of using fluorescence data as a proxy for cell density is that due to irregularities of gene regulation associated with stationary phase, fluorescence can continue to increase after bacteria in the population are no longer dividing (before stationary phase there is generally a linear relationship between OD600 optical density and fluorescence). To adjust for this, at each temperature we identify the time point at which bacteria are no longer dividing based on OD600 data. We use the fluorescence value at that time point in the fluorescence data set as the maximum value for fluorescence.

A.2.2.2. Data normalization

The relative fluorescence units (r.f.u.) of raw fluorescence data for the YFP molecule are on a scale from ~0-60,000 whereas the r.f.u. for the mCherry molecule are on a scale from ~0-5,000. We therefore employed a normalization protocol to convert the units of each to a common scale. To normalize, all raw fluorescence units were converted to the scale of r.f.u. from the *E. coli* 32 ˚C fluorescence data set and then divided by 10,000 to increase numeric compatibility with our fitting scripts.

A.3 Results

A.3.1. Estimation of intrinsic growth rates and competition coefficients from growth curves across temperatures using custom R scripts

To investigate the effect (if any) temperature has on the RFDs and NDs in our synthetic microbial bi-culture from chapter 2, we cultured *E. coli K12 MG1655* (expressing YFP) and *P. putida KT2440* (expressing mCh) in monoculture and co-culture across a range of temperatures from 32-40 ˚C (Figure A.1).

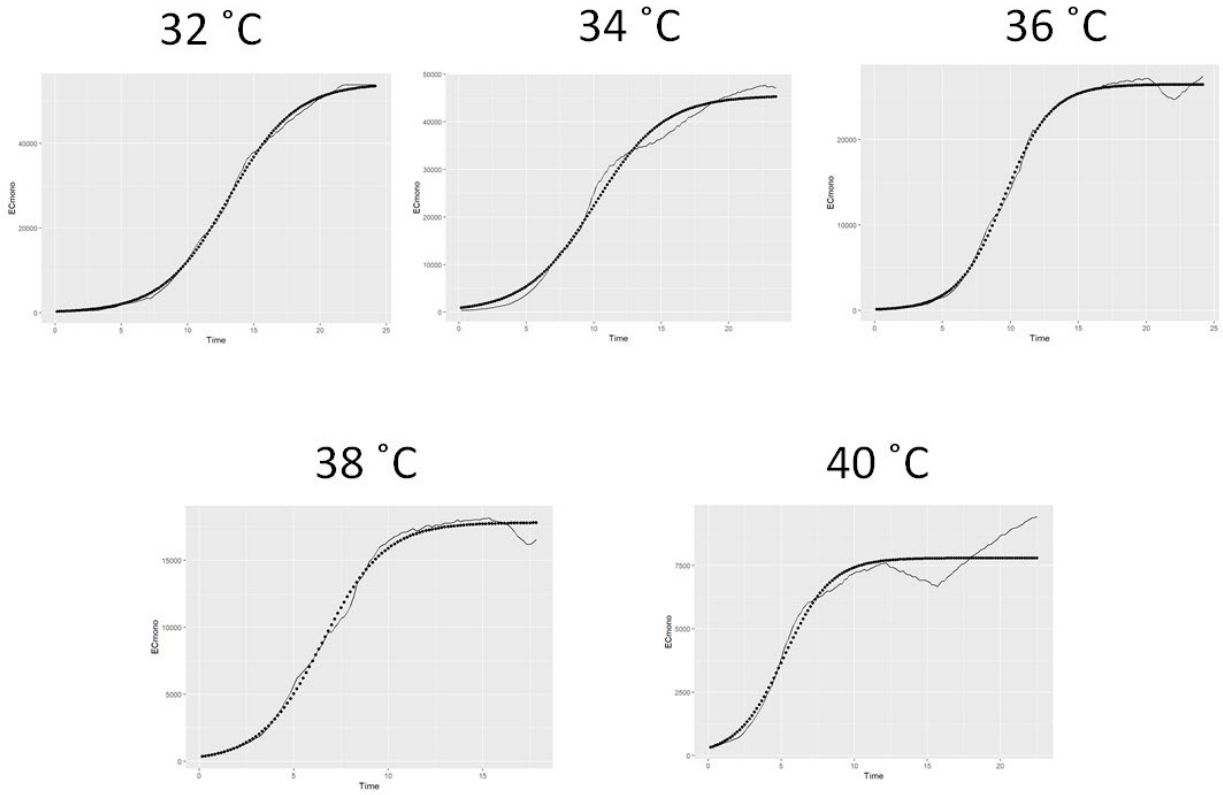


Figure A.1 Representative monoculture growth curves from E. coli. Raw data is shown in black lines and fitted curves are shown in black dots.

We used custom R scripts to fit the monoculture growth curves to a logistic growth differential equation:

$$\frac{dN_i}{dt} = r_i N_i (1 - \alpha_{ii} N_i)$$

$$\frac{dN_j}{dt} = r_j N_j \left(1 - \alpha_{jj} N_j\right)$$

where Ni represents the cell number of species i at a given time t, ri represents the μmax of species i and α_{ii} represents the intraspecific competition coefficient for species i (with the subscript $_{ii}$ notation indicating that competition coefficient α corresponds to the effect that species i has on itself). We used a linear regression with a least squares error minimization to estimate the intrinsic growth rates (ri and rj) at each temperature and to determine the intraspecific competition coefficients α_{ii} and α_{jj}.

Having determined estimates for the intrinsic growth rates and intraspecific competition coefficients for each species we next used those estimates as parameters in custom R scripts to fit the Lotka-Volterra model for continuous reproduction to co-culture growth curves:

$$\frac{dN_i}{dt} = r_i N_i\left(1 - \alpha_{ii}N_i - \alpha_{ij}N_j\right)$$

$$\frac{dN_j}{dt} = r_j N_j\left(1 - \alpha_{jj}N_j - \alpha_{ji}N_i\right)$$

where the interspecific competition coefficients α_{jj} and α_{ji} refer to the effect species j has on species and the effect species i has on species j, respectively). Again we used a linear

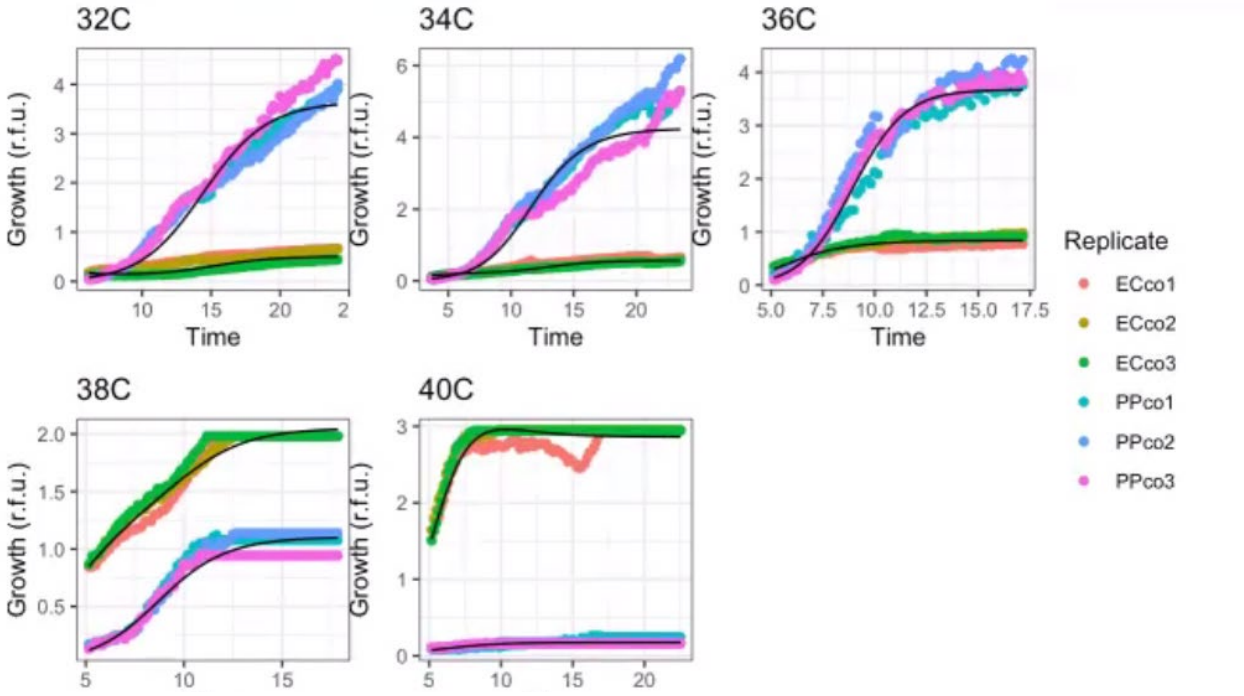


Figure A.2 Co-culture growth curves across temperatures with the Lotka-Volterra continuous reproduction model fits in black lines. E. coli is represented by the green, brown, and orange curves (replicates) whereas P. putida is represented by the pink, cyan, and blue curves (replicates).

regression with a least squares error minimization to fit the model to the data and generate parameter estimates.

Using this approach, the intra- and inter-specific competition coefficients estimates we generated can be found in Table 1.

Temperature (˚C)	α_{ii}	α_{jj}	α_{jj}	α_{ji}
32	7.98	0.41	-0.84	-0.94
34	4.48	0.40	-0.37	-1.25

36	1.32	0.18	-0.029	-0.4
38	0.8	1.18	-0.58	-0.15
40	0.25	6.25	1.57	-0.03

Table A.1 Intra- and inter-specific competition coefficient estimates from the R scripts.

A.3.2. Estimation of intrinsic growth rates and competition coefficients from growth curves across temperatures using difference equations and Excel Solver

As seen in Table 1, the intraspecific competition coefficients generally agree with the expected trends. Specifically, as temperature increases the intraspecific competition coefficients for *E. coli* decrease, whereas the intraspecific competition coefficients for *P. putida* increase, which agree with a qualitative analysis of the growth curves in Figure A.2. However, despite qualitatively good fits, 80% of the estimates generated for the interspecific competition coefficients are negative. The ecological interpretation of negative interspecific competition coefficients is that facilitation is occurring, where one species has a positive effect on the growth of the other. Not only is there no indication of facilitation from a qualitative analysis of the different between monoculture and co-culture growth curves for all species at all temperatures in this data set, but every single set of growths curves indicates the opposite (Figure A.3).

Although the fits our R scripts generate are qualitatively good, because they generate parameter estimates that do not make biological or ecological sense, we attempted to use a

different approach to reproduce competition coefficient estimates. We used a set of difference equations combined the with Excel Solver function to generate estimates for the intra- and interspecific competition coefficients (Figure A.4). Again, despite qualitatively excellent fits, the

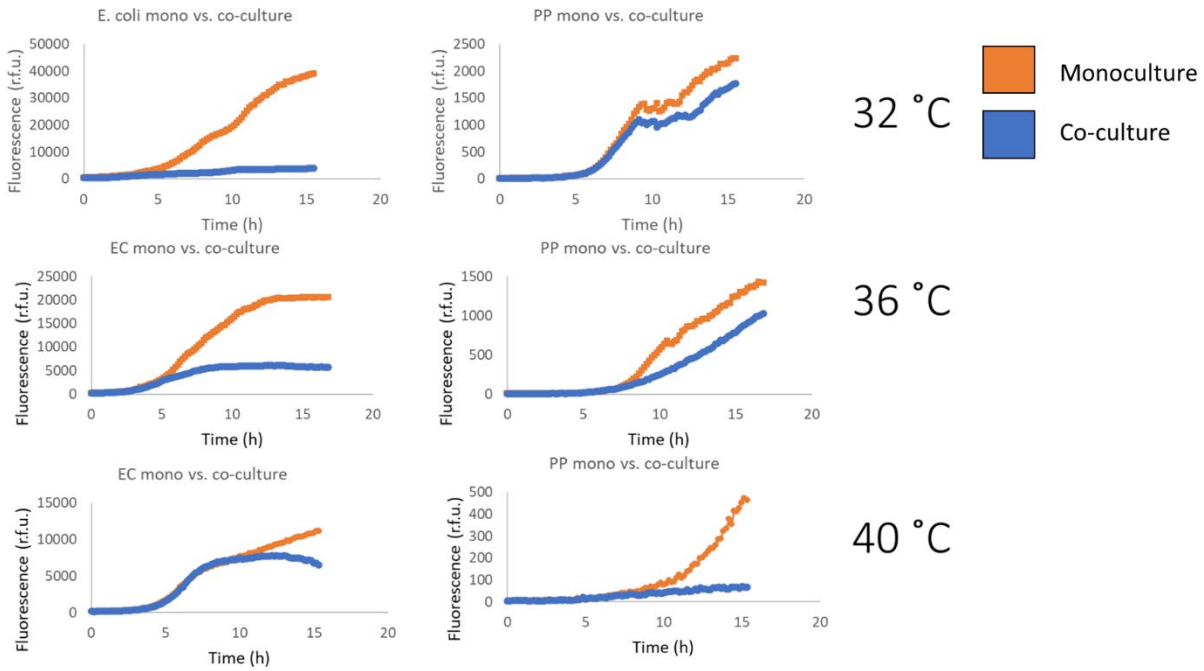


Figure A.3 Representative growth curves comparing monoculture to co-culture growth for E. coli and P. putida across temperatures. Across all temperatures co-culture growth qualitatively appears to be inhibited compared to monoculture growth.

several of the parameter estimates that are generated with the difference equations and Excel solver do not make biological or ecological sense (Table 2).

As can be seen in Table 2 of the parameter estimates generated using the difference equations and Solver in Excel, the intraspecific competition coefficient estimates agree in the direction of trends but not in magnitude with the estimates generated from the R script, but

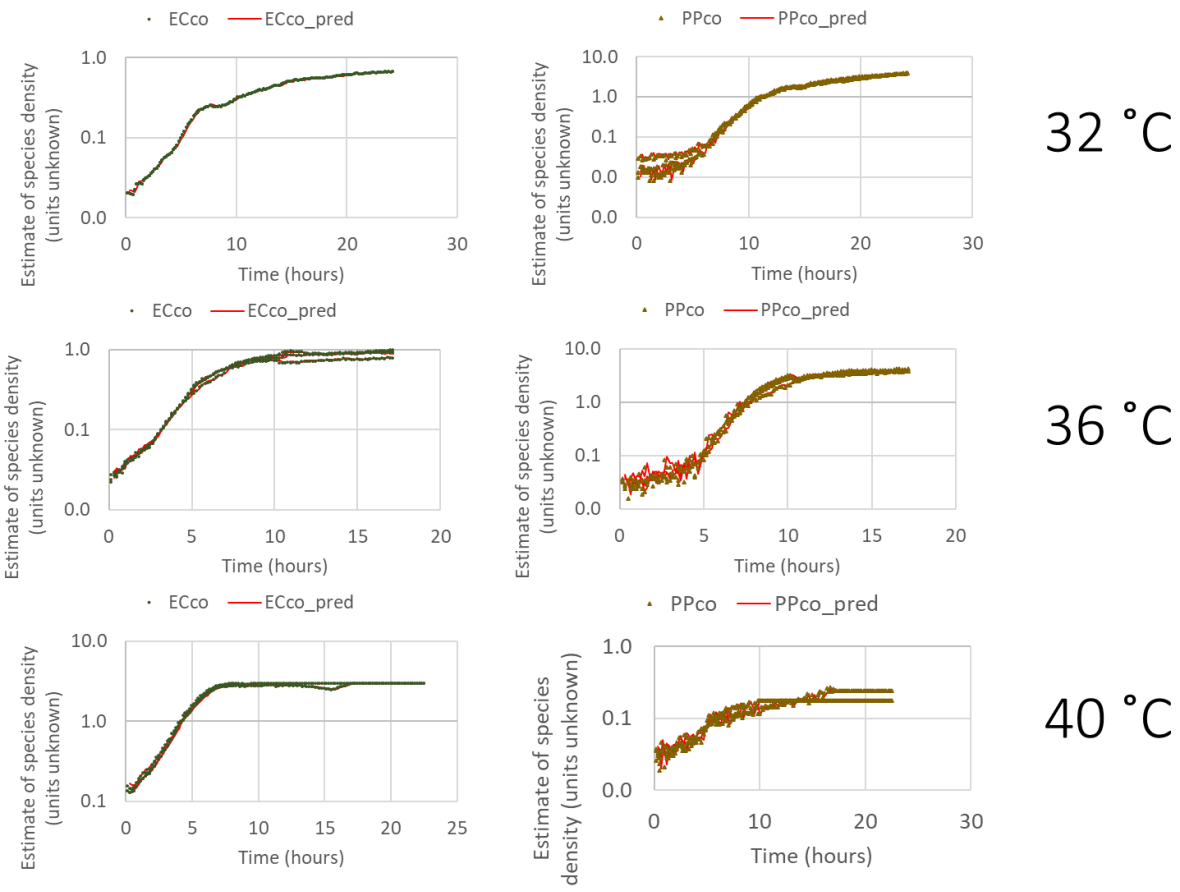


Figure A.4 Representative growth curves and the fits generated using difference equations in Excel. Data represented in dots and fits in red lines.

again a significant percentage (40%) of the interspecific competition coefficient estimates are negative, which as discussed previously disagrees with a qualitative analysis of growth curves (Figure A.3).

Temperature (˚C)	α_{ii}	α_{jj}	α_{jj}	α_{ji}
32	12.12	0.15	-1.06	6.94
34	8.68	0.38	-0.11	4.78
36	4.22	0.81	-0.14	1.09

38	3.32	2.65	-1.76	0.09
40	1.45	14.76	0.62	0.17

Table A.2 Intra- and inter-specific competition coefficient estimates generated using difference equations and Solver.

Because the interspecific competition coefficients estimates we have generated with two orthogonal approaches produce negative interspecific competition coefficients, which contradict a qualitative analysis of growth curve data (Tables 1-2, Figure A.3) we do not feel it is appropriate to use the competition coefficient parameter estimates to generate or discuss RFD and ND estimates. Additionally, as the equations to generate RFDs and NDs require taking the square root of a ratio of competition coefficient products:

$$RFD = \sqrt{\frac{\alpha_{ij}\alpha_{ii}}{\alpha_{jj}\alpha_{ji}}}$$

$$ND = 1 - \sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}}$$

in many cases negative competition coefficients result in RFD and ND estimates that are not real numbers.

A.4. Discussion

In Chapter 2 we demonstrated a novel technique to enable coexistence and program relative species abundance in synthetic microbial communities by regulating temperature. However, aspects of the ways in which community population dynamics responded to temperature were unexpected and intriguing. Due to the dearth of investigation by the synthetic biology and synthetic ecology communities into this type of phenomena, we turned to

over a century of research performed by natural ecologists on mechanisms of species coexistence in naturally occurring communities.

Modern coexistence theory has provided a broadly applicable, mathematically rigorous framework to probe coexistence mechanisms across a wide variety of communities displaying various categories of species interactions. This approach uses community population dynamics data in monoculture and co-culture to estimate two parameters, relative fitness differences (RFDs) and niche differences (NDs). While a number of empirical methods are available to generate RFD and ND estimates from experimental data, we chose the method based on the Lotka-Volterra model for continuous reproduction because it is most amenable to our experimental design. We used two different workflows to implement this approach with our experimental monoculture and co-culture data from the synthetic microbial system we developed in Chapter 2. However, both of these approaches generated estimates of interspecific competition coefficients that do not make sense within a biological or ecological context. Specifically, many of the interspecific competition coefficient parameter estimates we generated are negative, which suggests facilitation between species, and a qualitative analysis of the difference between monoculture and co-culture growth curves at almost every temperature indicate an inhibitory as opposed to facilitative interaction. In light of this, it is important to consider possible explanations that might explain the discrepancy between our quantitative and qualitative analyses.

One potential explanation is that our implementation of the model is flawed. It is possible that either bugs in our R-code or a mistake in the way we have input the model equations is resulting in some systemic error. However, the fact that two independent

approaches for parameter estimation generate intraspecific competition coefficients with trends that agree, and that those trends agree with what we would expect for the intraspecific competition coefficient trends suggests in my opinion that the negative values generated for the interspecific competition coefficients are not due to a systemic error. There is no systemic error I can think of that would affect the interspecific but not the intraspecific competition coefficients in our two different parameter estimation techniques.

Assuming there is no systemic error in our work, another possible explanation for this discrepancy is that our system does not meet one or more of the assumptions for Chesson's model. For example, one assumption that must be met for Chesson's theory to apply is that the products of the intraspecific competition coefficients must be greater than the interspecific competition coefficients. While this is sometimes but not always the case for the competition coefficients we have generated, if this assumption is not met it would mean the competition coefficients cannot be used to calculate accurate RFDs and NDs, but it would not necessarily explain our inability to generate competition coefficients that agree with our qualitative expectations.

In light of this line of thinking, it is also important to consider whether our system meets the assumptions of the Lotka-Volterra interspecific competition model. One crucial assumption for the Lotka-Volterra model is that the carrying capacities and competition coefficients for both species be constants (Beals et al., 1999; Begon, Harper, & Townsend, 1996). In retrospect, it is entirely possible that this assumption is not met within our system. Because we do not know the specific interspecific competition mechanisms at work in our system (and indeed, it is one of the strengths of Chesson's model that one need not know the precise competitive

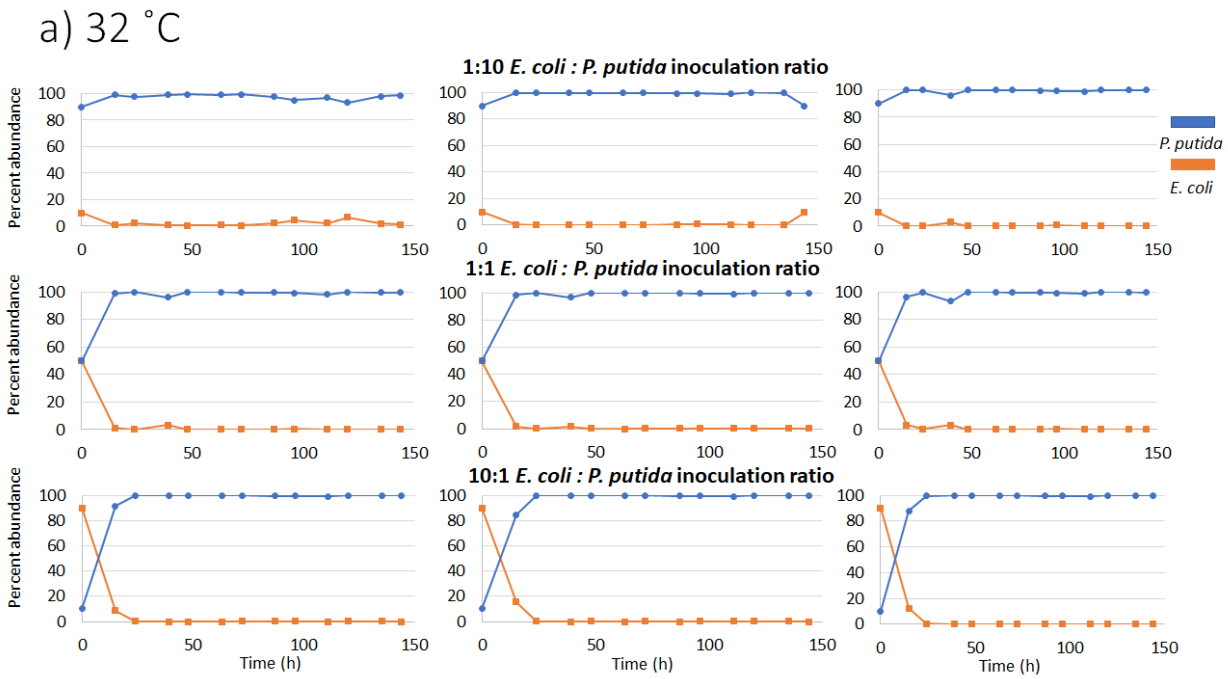
mechanisms), we certainly cannot rule out the possibility that the strength of the competition changes with respect to parameters such as population density. So, the prospect that the interspecific competition coefficients in our system are not constant is one possible explanation for the dissonance between our quantitative and qualitative analyses.

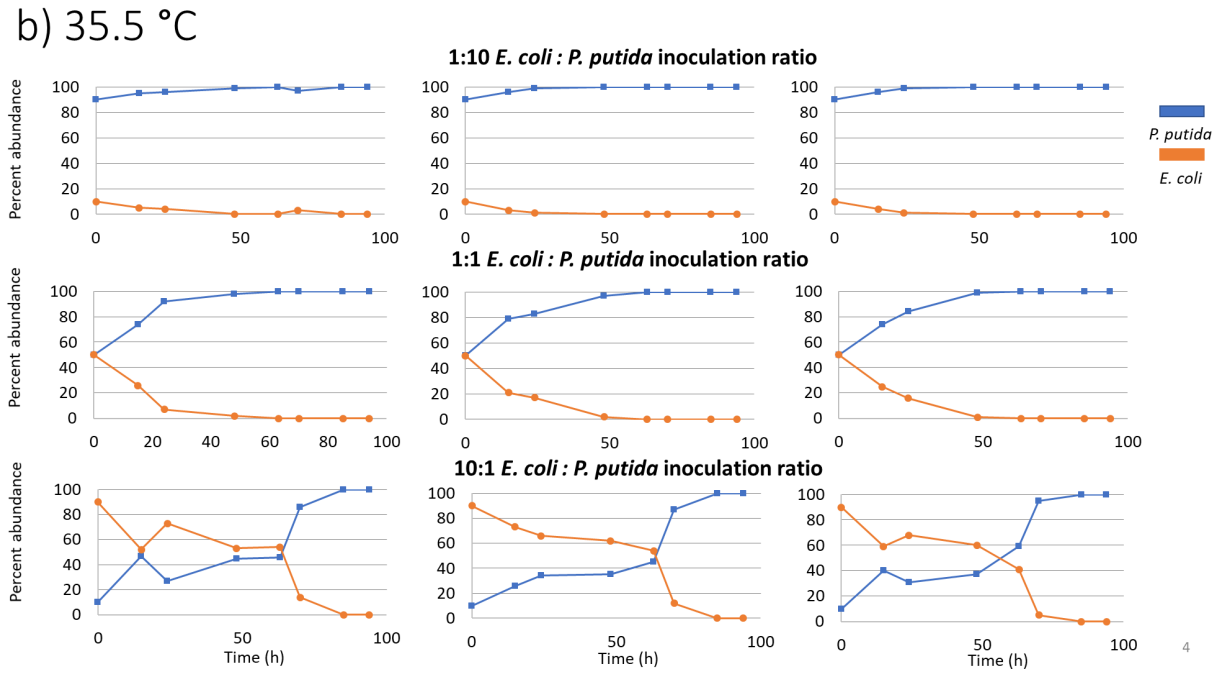
Due to our failure thus far to generate parameter estimates that pass a "sense check", we cannot yet meaningfully discuss the ways in which (if at all) RFDs and NDs change in response to temperature. However, it is still possible to use a qualitative analysis of growth curve data to hypothesize what such a relationship might look like (if it exists). As discussed above, the trend we observe in the intraspecific competition coefficients using parameter estimates from both our R script and the Excel Solver approach make sense in that as species growth is increasingly negatively affected by temperature, their intraspecific competition coefficient increases. It was our initial hypothesis that a trend would also be observed in the interspecific competition coefficients. Specifically, as temperature increases, the interspecific competition coefficient for the effect of *P. putida* on *E. coli* (α_{jj}) would decrease and the interspecific competition coefficient for the effect of *E. coli* on *P. putida* (α_{ji}) would increase. It is also tempting to predict that there might be some type of parabolic relationship between temperature and RFDs. At relative temperature extremes (e.g. 32 and 40 ˚C) it is possible that RFDs would be highest, whereas as the temperature increases from 32 ˚C the RFDs would be expected to decrease and then begin to increase again. It is not immediately clear whether NDs will be affected in any way by temperature. If significant shifts in ND are observed in response to temperature, one possible explanation is that the activity of specific proteins, enzymes, receptors, and transporters related to resource and nutrients uptake and utilization may

change with temperature, which could potentially affect niche differences. However, without reliable parameter estimates, it is impossible to discuss the effects of temperature on RFDs and NDs with any certainty. We look forward to exploring new ways to optimize our models to generate parameter estimates that make sense within biological and ecological contexts.

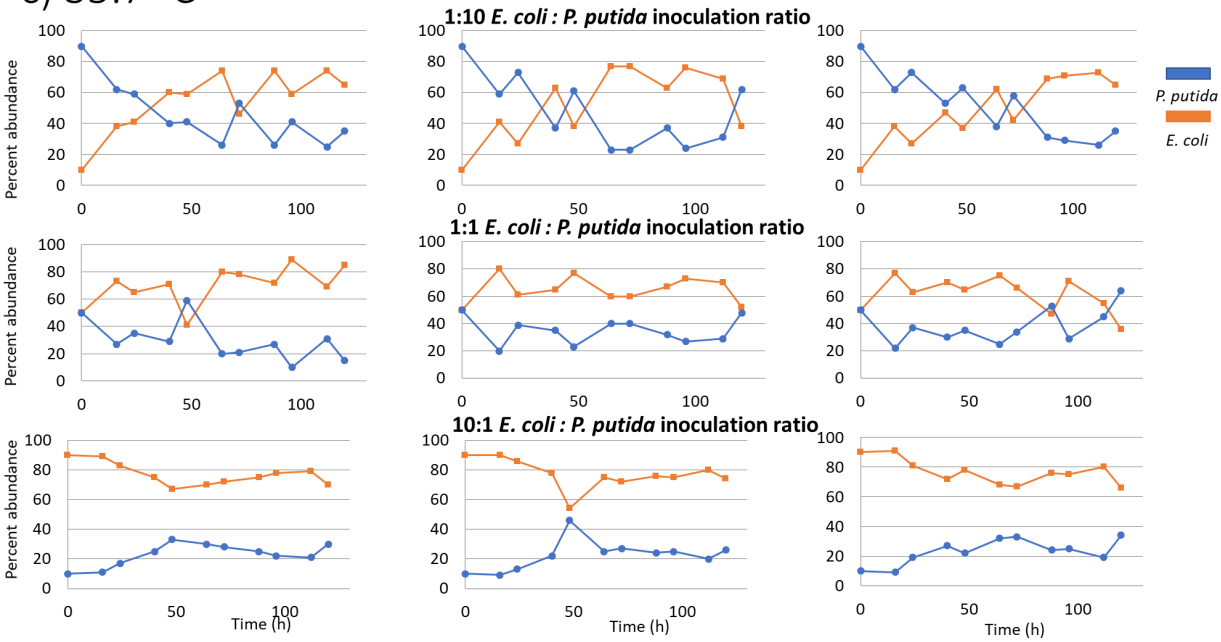
Appendix B

Time series data from all constant temperature regime experiments

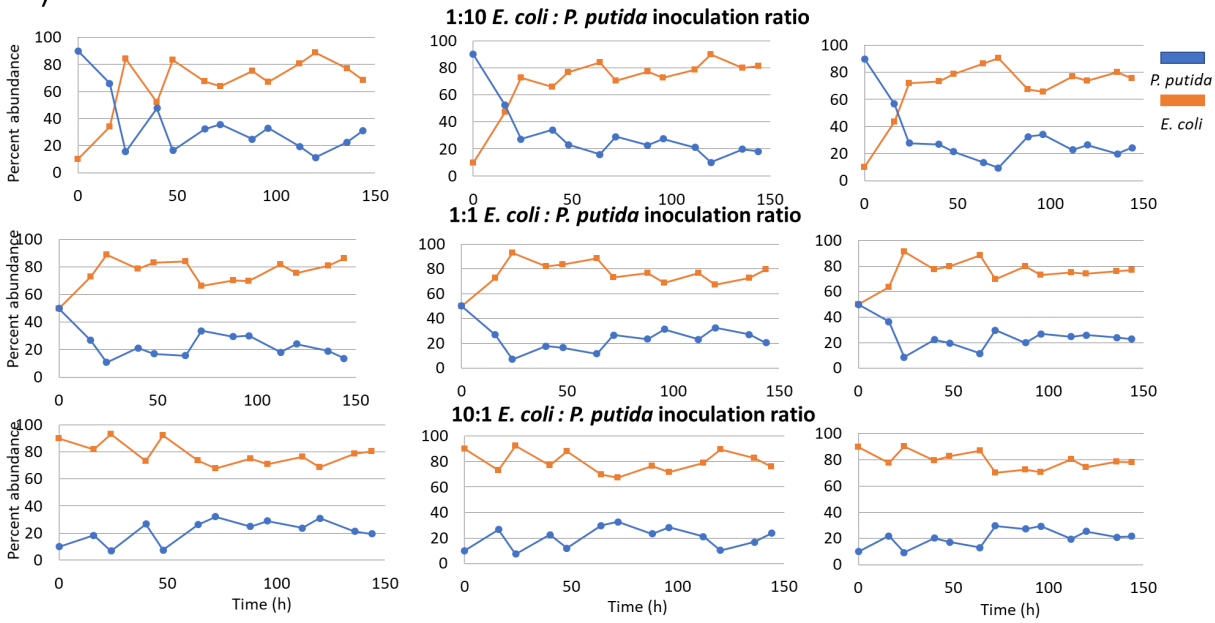




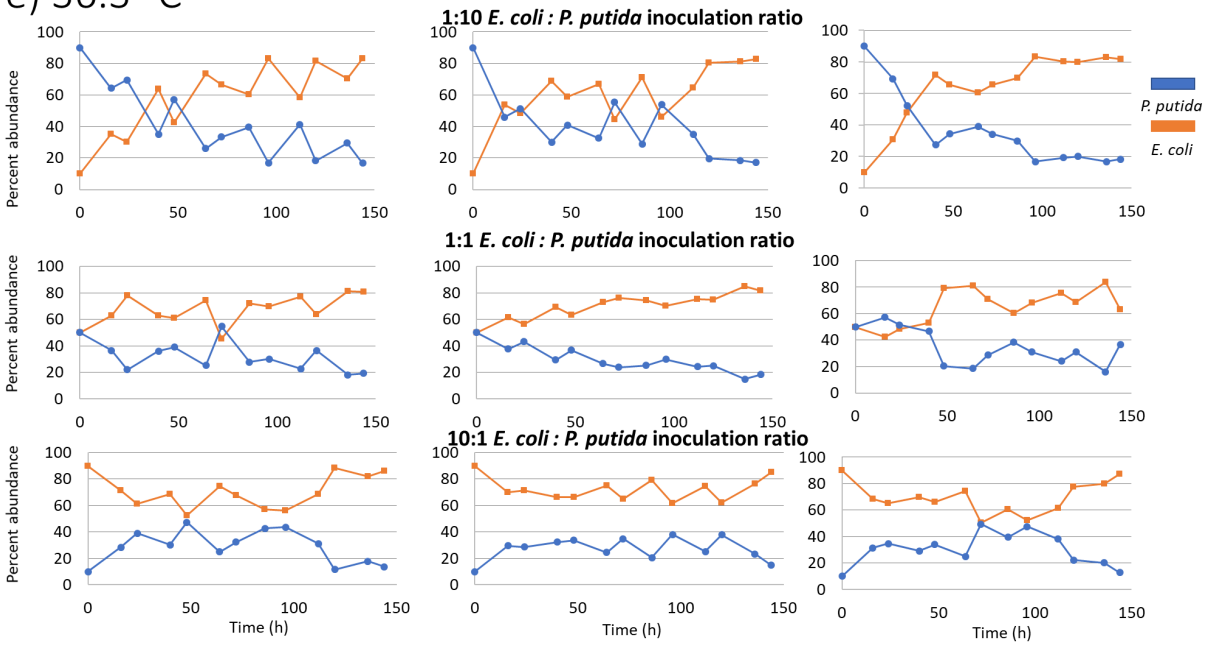
c) 35.7 °C


1:10 E. coli : P. putida inoculation ratio
1:1 E. coli : P. putida inoculation ratio
10:1 E. coli : P. putida inoculation ratio
Percent abundance
Time (h)
P. putida
E. coli

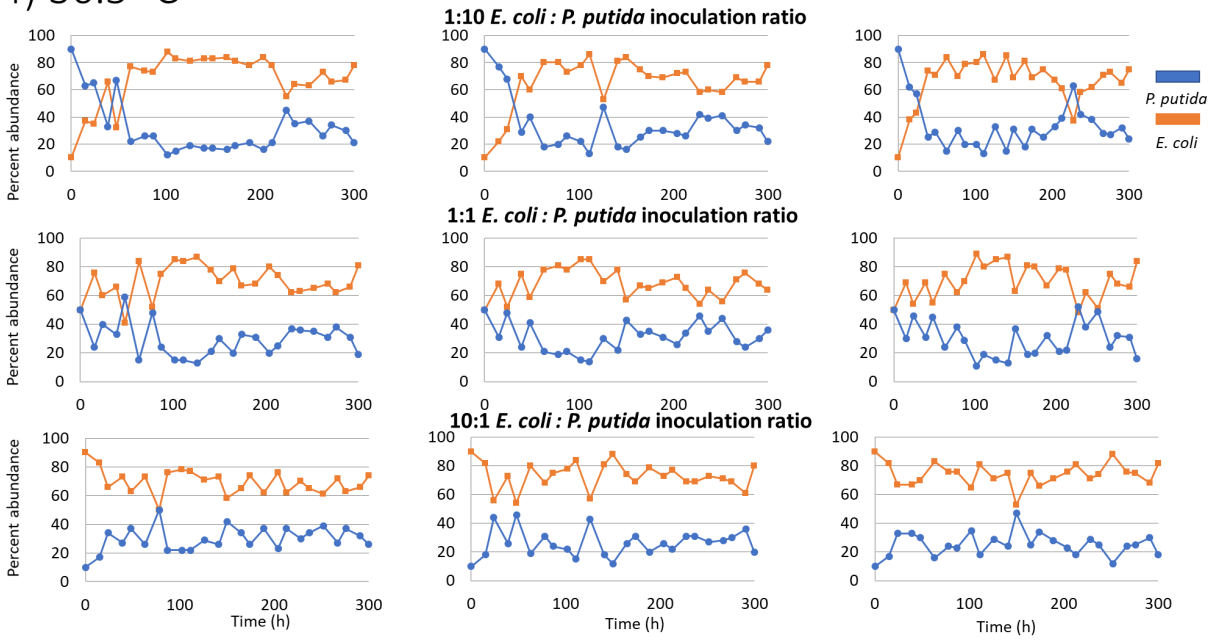
d) 36.1 °C


1:10 E. coli : P. putida inoculation ratio
1:1 E. coli : P. putida inoculation ratio
10:1 E. coli : P. putida inoculation ratio
Percent abundance
Time (h)
P. putida
E. coli

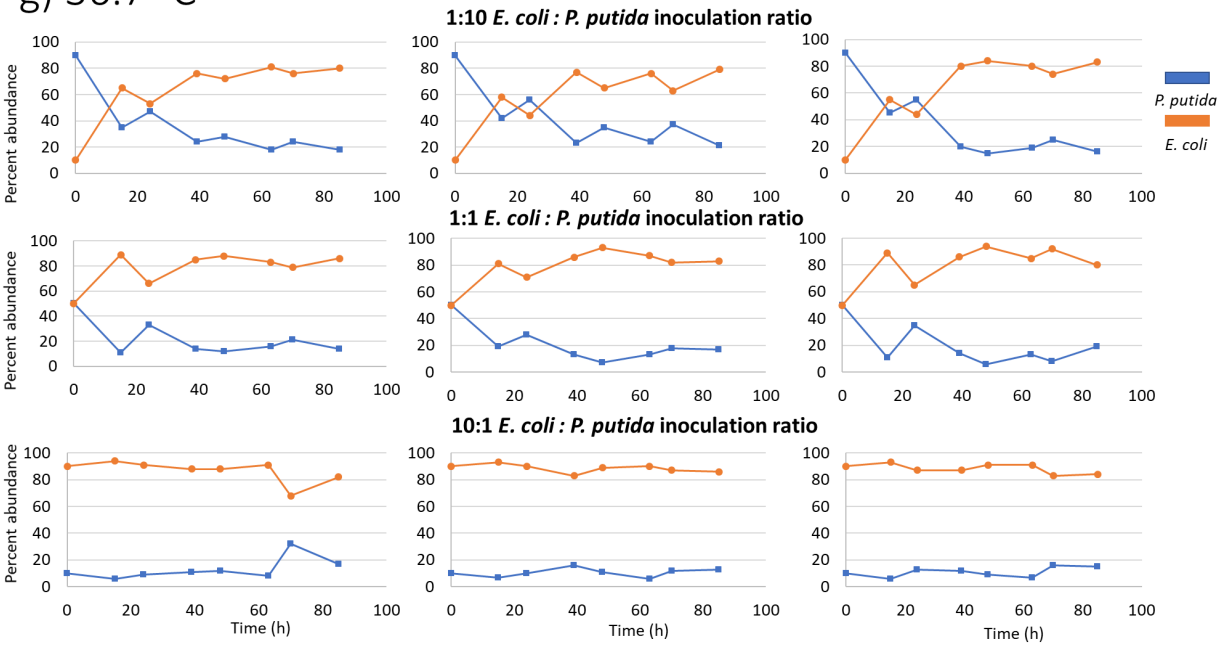
e) 36.3 °C


1:10 E. coli : P. putida inoculation ratio
1:1 E. coli : P. putida inoculation ratio
10:1 E. coli : P. putida inoculation ratio
Percent abundance
Time (h)
P. putida
E. coli

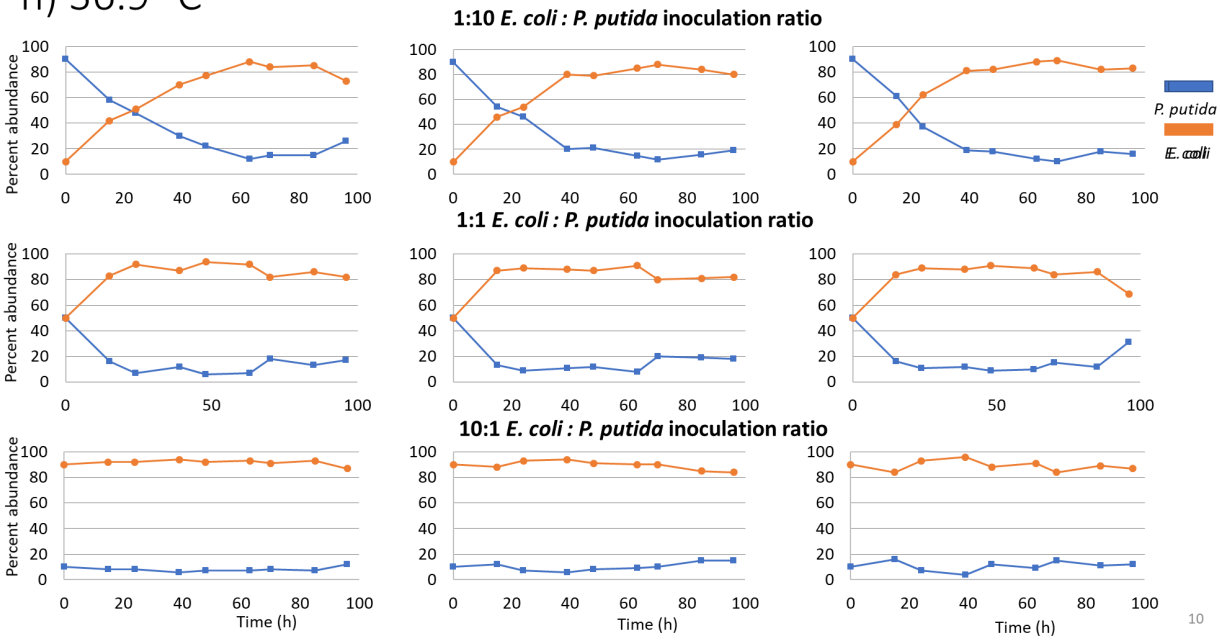
f) 36.5 °C


1:10 E. coli : P. putida inoculation ratio
1:1 E. coli : P. putida inoculation ratio
10:1 E. coli : P. putida inoculation ratio
Percent abundance
Time (h)
P. putida
E. coli

g) 36.7 ˚C



h) 36.9 °C



10

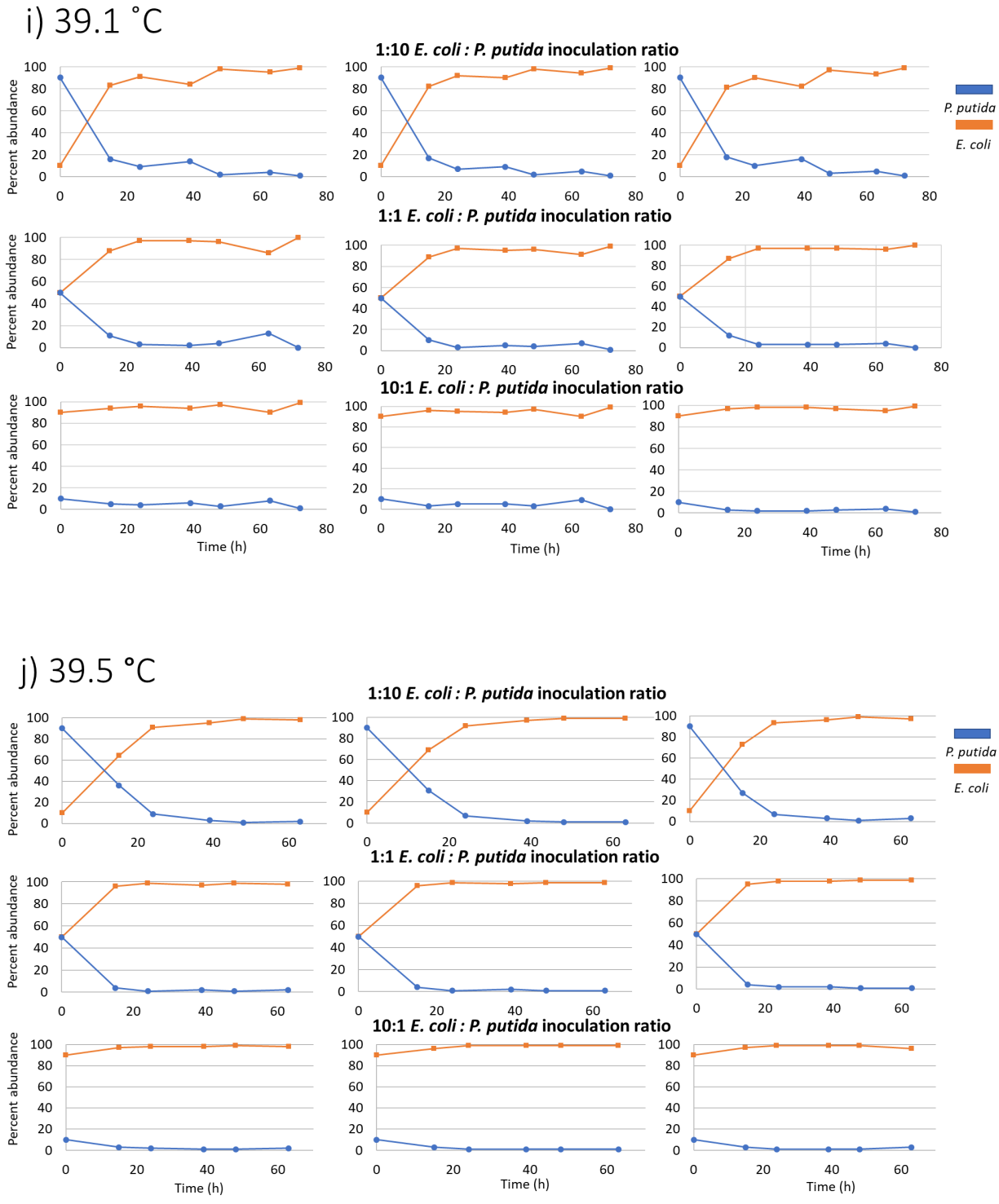


Figure B.1 Time series data from all constant temperature regime experiments.

Appendix C

Representative confocal microscopy and flow cytometry data from fluorescent strains

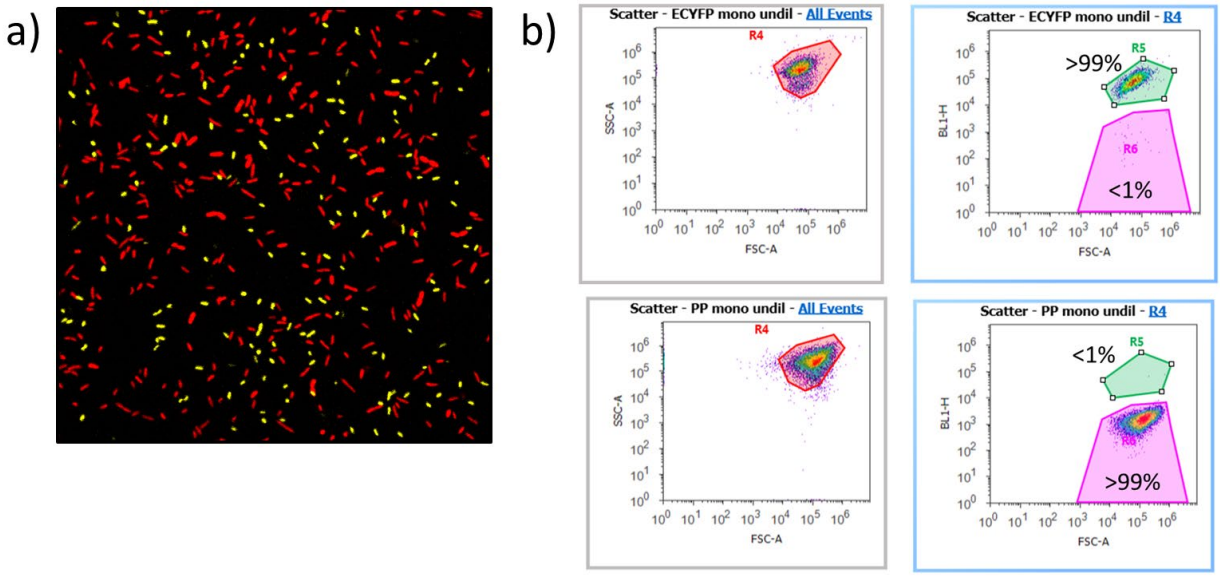


Figure C.1 a) Confocal microscopy image of co-culture. b) Constitutive expression of a chromosomally integrated cassette for Yellow fluorescent protein (E. coli) and mCherry (P. putida) were used quantify the community composition of bi-cultures via flow cytometry.

Appendix D

Representative temperature fluctuation data

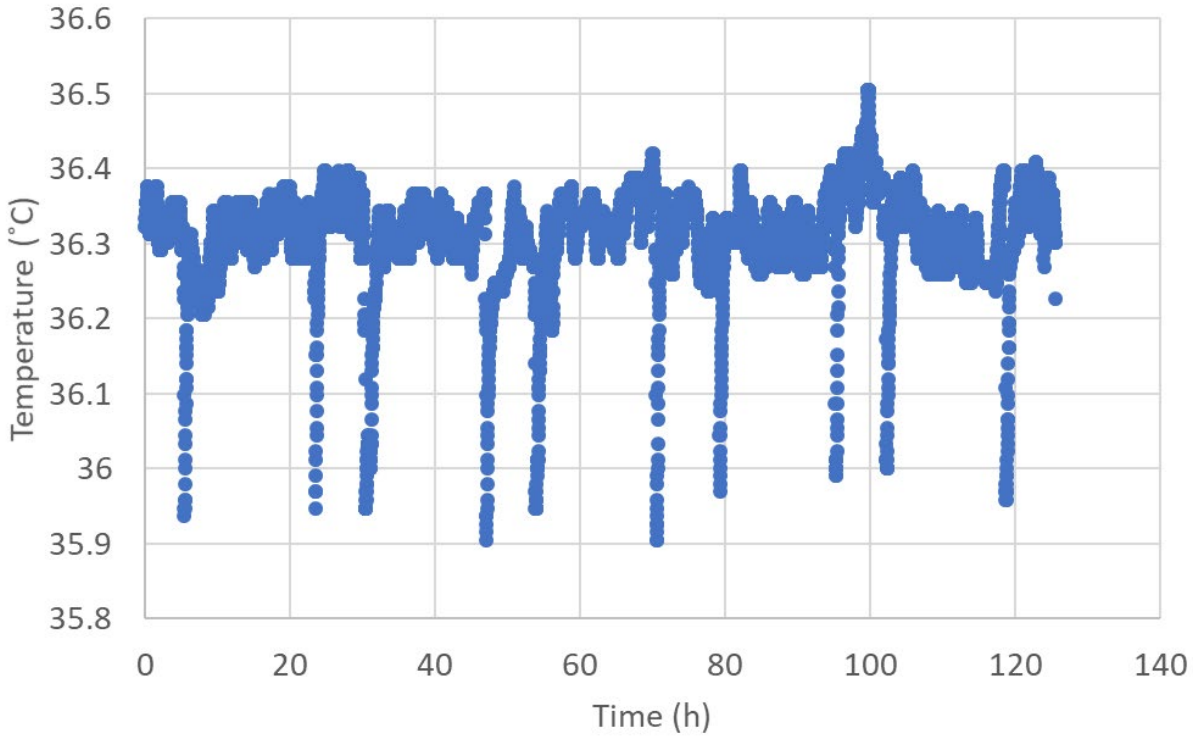


Figure D.1 Representative graph of temperatures over time as recorded inside the incubator when the incubator was set to 36.3 ˚C.

Bibliography

Aoki, M., Ehara, M., Saito, Y., Yoshioka, H., Miyazaki, M., Saito, Y., . . . Imachi, H. (2014). A Long-Term Cultivation of an Anaerobic Methane-Oxidizing Microbial Community from Deep-Sea Methane-Seep Sediment Using a Continuous-Flow Bioreactor. *PLoS One, 9*(8), e105356. doi:10.1371/journal.pone.0105356

Barrangou, R. (2015). The roles of CRISPR–Cas systems in adaptive immunity and beyond. *Current Opinion in Immunology, 32*, 36-41. doi:https://doi.org/10.1016/j.coi.2014.12.008

Beals, M., Gross, L., & Harrell, S. (1999). Interspecific competition: Lotka-Volterra. Retrieved from http://www.tiem.utk.edu/~gross/bioed/bealsmodules/competition.html

Begon, M., Harper, J. L., & Townsend, C. R. (1996). *Ecology: Individuals, Populations, and Communities* (3rd ed.). Cambridge, MA: Blackwell Science Ltd.

Belda, E., van Heck, R. G., Jose Lopez-Sanchez, M., Cruveiller, S., Barbe, V., Fraser, C., . . . Medigue, C. (2016). The revisited genome of Pseudomonas putida KT2440 enlightens its value as a robust metabolic chassis. *Environ Microbiol, 18*(10), 3403-3424. doi:10.1111/1462-2920.13230

Bennett, A. F., Lenski, R. E., & Mittler, J. E. (1992). Evolutionary Adaptation to Temperature. I. Fitness Responses of Escherichia coli to Changes in its Thermal Environment. *Evolution, 46*(1), 16-30. doi:10.2307/2409801

Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., . . . Shao, Y. (1997). The complete genome sequence of Escherichia coli K-12. *Science, 277*(5331), 1453-1462. doi:10.1126/science.277.5331.1453

Bouwer, E. J., & Zehnder, A. J. B. (1993). Bioremediation of organic compounds — putting microbial metabolism to work. *Trends in Biotechnology, 11*(8), 360-367. doi:https://doi.org/10.1016/0167-7799(93)90159-7

Brenner, K., You, L., & Arnold, F. H. (2008). Engineering microbial consortia: a new frontier in synthetic biology. *Trends Biotechnol, 26*(9), 483-489. doi:10.1016/j.tibtech.2008.05.004

Britannica, E. (2020). Engineering.

Broger, T., Odermatt, R. P., Huber, P., & Sonnleitner, B. (2011). Real-time on-line flow cytometry for bioprocess monitoring. *J Biotechnol, 154*(4), 240-247. doi:10.1016/j.jbiotec.2011.05.003

Burgess-Brown, N. A., Sharma, S., Sobott, F., Loenarz, C., Oppermann, U., & Gileadi, O. (2008). Codon optimization can improve expression of human genes in Escherichia coli: A multi-gene study. *Protein Expression and Purification, 59*(1), 94-102. doi:https://doi.org/10.1016/j.pep.2008.01.008

Butala, M., Zgur-Bertok, D., & Busby, S. J. (2009). The bacterial LexA transcriptional repressor. *Cell Mol Life Sci, 66*(1), 82-93. doi:10.1007/s00018-008-8378-6

Butterfield, G. L., Lajoie, M. J., Gustafson, H. H., Sellers, D. L., Nattermann, U., Ellis, D., . . . Baker, D. (2017). Evolution of a designed protein assembly encapsulating its own RNA genome. *Nature, 552*(7685), 415-420. doi:10.1038/nature25157

Cai, Q., Qiao, L., Wang, M., He, B., Lin, F. M., Palmquist, J., . . . Jin, H. (2018). Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes. *Science, 360*(6393), 1126-1129. doi:10.1126/science.aar4142

Campbell, A. (2003). The future of bacteriophage biology. *Nature Reviews Genetics, 4*(6), 471-477. doi:10.1038/nrg1089

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., . . . Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature, 486*(7401), 59-67. doi:10.1038/nature11148

Carroll, I. T., Cardinale, B. J., & Nisbet, R. M. (2011). Niche and fitness differences relate the maintenance of diversity to ecosystem function. *Ecology, 92*(5), 1157-1165. doi:10.1890/10-0302.1

Casjens, S. R., & Hendrix, R. W. (2015). Bacteriophage lambda: Early pioneer and still relevant. *Virology, 479-480*, 310-330. doi:10.1016/j.virol.2015.02.010

CBInsights. (2019, 01/04/2019). From Alibab to Zynga: 40 of the best VC bets of all time and what we can learn from them. *Research briefs.* Retrieved from https://www.cbinsights.com/research/best-venture-capital-investments/

Charbonneau, M. R., Isabella, V. M., Li, N., & Kurtz, C. B. (2020). Developing a new class of engineered live bacterial therapeutics to treat human diseases. *Nature Communications, 11*(1), 1738. doi:10.1038/s41467-020-15508-1

Chen, K., Liu, Y., & Chen, Q. (2009). *HOW MUCH DOES HISTORY MATTER? AN ANALYSIS OF THE GEOGRAPHIC DISTRIBUTION OF VENTURE CAPITAL INVESTMENT IN THE U.S. BIOTECHNOLOGY INDUSTRY*.

Chesson, P. (2000). Mechanisms of Maintenance of Species Diversity. *Annual Review of Ecology and Systematics, 31*(1), 343-366. doi:10.1146/annurev.ecolsys.31.1.343

Chiang, Y. N., Penadés, J. R., & Chen, J. (2019). Genetic transduction by phages and chromosomal islands: The new and noncanonical. *PLoS pathogens, 15*(8), e1007878-e1007878. doi:10.1371/journal.ppat.1007878

Council, N. R. (2014). *Technological Challenges in Antibiotic Discovery and Development: A Workshop Summary*. Washington, DC: The National Academies Press.

Darwin, C. (1859). *On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life*: London : John Murray, 1859.

Davison, B. H., & Stephanopoulos, G. (1986). Effect of pH oscillations on a competing mixed culture. *Biotechnol Bioeng, 28*(8), 1127-1137. doi:10.1002/bit.260280802

Deatherage, B. L., & Cookson, B. T. (2012). Membrane vesicle release in bacteria, eukaryotes, and archaea: a conserved yet underappreciated aspect of microbial life. *Infect Immun, 80*(6), 1948-1957. doi:10.1128/IAI.06014-11

DeLisa, M. P., Valdes, J. J., & Bentley, W. E. (2001). Quorum signaling via AI-2 communicates the "Metabolic Burden" associated with heterologous protein production in Escherichia coli. *Biotechnol Bioeng, 75*(4), 439-450. doi:10.1002/bit.10034

Dunn, A. K., & Handelsman, J. (2002). Toward an understanding of microbial communities through analysis of communication networks. *Antonie van Leeuwenhoek, 81*(1), 565-574. doi:10.1023/A:1020565807627

Dutta, S., Whicher, J. R., Hansen, D. A., Hale, W. A., Chemler, J. A., Congdon, G. R., . . . Skiniotis, G. (2014). Structure of a modular polyketide synthase. *Nature, 510*(7506), 512-517. doi:10.1038/nature13423

Eggers, C. H., Gray, C. M., Preisig, A. M., Glenn, D. M., Pereira, J., Ayers, R. W., . . . Moeller, J. T. (2016). Phage-mediated horizontal gene transfer of both prophage and heterologous DNA by φBB-1, a bacteriophage of Borrelia burgdorferi. *Pathogens and Disease, 74*(9). doi:10.1093/femspd/ftw107

ExPASy. *ViralZone.* Retrieved from https://viralzone.expasy.org/512?outline=all_by_species

Farzadfard, F., & Lu, T. K. (2014). Synthetic biology. Genomically encoded analog memory with precise in vivo DNA writing in living cell populations. *Science (New York, N.Y.), 346*(6211), 1256272-1256272. doi:10.1126/science.1256272

Ferguson, B. J., Indrasumunar, A., Hayashi, S., Lin, M. H., Lin, Y. H., Reid, D. E., & Gresshoff, P. M. (2010). Molecular analysis of legume nodule development and autoregulation. *J Integr Plant Biol, 52*(1), 61-76. doi:10.1111/j.1744-7909.2010.00899.x

Forbes, S., Dobson, C. B., Humphreys, G. J., & McBain, A. J. (2014). Transient and sustained bacterial adaptation following repeated sublethal exposure to microbicides and a novel human antimicrobial peptide. *Antimicrobial agents and chemotherapy, 58*(10), 5809-5817. doi:10.1128/AAC.03364-14

Franklin, R. E., & Gosling, R. G. (1953). Molecular Configuration in Sodium Thymonucleate. *Nature, 171*(4356), 740-741. doi:10.1038/171740a0

Fridman, O., Goldberg, A., Ronin, I., Shoresh, N., & Balaban, N. Q. (2014). Optimization of lag time underlies antibiotic tolerance in evolved bacterial populations. *Nature, 513*(7518), 418-421. doi:10.1038/nature13469

Galkin, V. E., Yu, X., Bielnicki, J., Ndjonka, D., Bell, C. E., & Egelman, E. H. (2009). Cleavage of bacteriophage lambda cI repressor involves the RecA C-terminal domain. *J Mol Biol, 385*(3), 779-787. doi:10.1016/j.jmb.2008.10.081

Gause, G. F. (1934). EXPERIMENTAL ANALYSIS OF VITO VOLTERRA'S MATHEMATICAL THEORY OF THE STRUGGLE FOR EXISTENCE. *Science, 79*(2036), 16-17. doi:10.1126/science.79.2036.16-a

Ghoul, M., & Mitri, S. (2016). The Ecology and Evolution of Microbial Competition. *Trends Microbiol, 24*(10), 833-845. doi:10.1016/j.tim.2016.06.011

Glick, B. R. (1995). Metabolic load and heterologous gene expression. *Biotechnology Advances, 13*(2), 247-261. doi:https://doi.org/10.1016/0734-9750(95)00004-A

Godwin, C. M., Chang, F.-H., & Cardinale, B. J. (2020). An empiricist's guide to modern coexistence theory for competitive communities. *Oikos, n/a*(n/a). doi:10.1111/oik.06957

Goeddel, D. V., Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., . . . Riggs, A. D. (1979). Expression in Escherichia coli of chemically synthesized genes for human insulin. *Proceedings of the National Academy of Sciences, 76*(1), 106. doi:10.1073/pnas.76.1.106

Greene, J. J., & Rao, V. B. (1998). *Recombinant DNA principles and methodologies* (Vol. New York :). New York :: Marcel Dekker.

Grinnell, J. (1904). The origin and distribution of the chestnut-backed chickadee. *Auk, 21*(3), 364-382.

Guan, J., Ibarra, D., & Zeng, L. (2019). The role of side tail fibers during the infection cycle of phage lambda. *Virology, 527*, 57-63. doi:https://doi.org/10.1016/j.virol.2018.11.005

Guynet, C., Le, P. T. N., Chandler, M., & Ton-Hoang, B. (2020). Detection and Characterization of Transposons in Bacteria. In F. de la Cruz (Ed.), *Horizontal Gene Transfer: Methods and Protocols* (pp. 81-90). New York, NY: Springer US.

Harder, W., & Veldkamp, H. (1971). Competition of marine psychrophilic bacteria at low temperatures. *Antonie van Leeuwenhoek, 37*(1), 51-63. doi:10.1007/BF02218466

Hardin, G. (1960). The competitive exclusion principle. *Science, 131*(3409), 1292-1297. doi:10.1126/science.131.3409.1292

Hendrix, R. W. (1983). *Lambda II*: Cold Spring Harbor Laboratory.

Hershey, A. D. (1971). *The Bacteriophage lambda*, [Cold Spring Harbor, N.Y.

Höffner, K., & Barton, P. I. (2014). Design of Microbial Consortia for Industrial Biotechnology. In *Proceedings of the 8th International Conference on Foundations of Computer-Aided Process Design* (pp. 65-74).

Hohn, B. (1975). DNA as substrate for packaging into bacteriophage lambda, in vitro. *Journal of Molecular Biology, 98*(1), 93-106. doi:https://doi.org/10.1016/S0022-2836(75)80103-3

Hohn, B. (1983). DNA sequences necessary for packaging of bacteriophage lambda DNA. *Proceedings of the National Academy of Sciences, 80*(24), 7456. doi:10.1073/pnas.80.24.7456

Hohn, B., & Murray, K. (1977). Packaging recombinant DNA molecules into bacteriophage particles in vitro. *Proceedings of the National Academy of Sciences, 74*(8), 3259. doi:10.1073/pnas.74.8.3259

Homayun, B., Lin, X., & Choi, H.-J. (2019). Challenges and Recent Progress in Oral Drug Delivery Systems for Biopharmaceuticals. *Pharmaceutics, 11*(3), 129. Retrieved from https://www.mdpi.com/1999-4923/11/3/129

Hooven, M. D. (2017, 12/22/2017). Opportunities and challenges in biologic drug delivery. Retrieved from https://www.americanpharmaceuticalreview.com/Featured-Articles/345540-Opportunities-and-Challenges-in-Biologic-Drug-Delivery/

Houde, A., Kademi, A., & Leblanc, D. (2004). Lipases and their industrial applications. *Applied Biochemistry and Biotechnology, 118*(1), 155-170. doi:10.1385/ABAB:118:1-3:155

Hsu, B. B., Gibson, T. E., Yeliseyev, V., Liu, Q., Lyon, L., Bry, L., . . . Gerber, G. K. (2019). Dynamic Modulation of the Gut Microbiota and Metabolome by Bacteriophages in a Mouse Model. *Cell Host Microbe, 25*(6), 803-814 e805. doi:10.1016/j.chom.2019.05.001

Hsu, B. B., Way, J. C., & Silver, P. A. (2020). Stable Neutralization of a Virulence Factor in Bacteria Using Temperate Phage in the Mammalian Gut. *mSystems, 5*(1), e00013-00020. doi:10.1128/mSystems.00013-20

Humphries, J., Xiong, L., Liu, J., Prindle, A., Yuan, F., Arjes, H. A., . . . Süel, G. M. (2017). Species-Independent Attraction to Biofilms through Electrical Signaling. *Cell, 168*(1-2), 200-209.e212. doi:10.1016/j.cell.2016.12.014

Hutchinson, G. E. (1961). The Paradox of the Plankton. *The American Naturalist, 95*(882), 137-145. Retrieved from www.jstor.org/stable/2458386

Inamdar, M. M., Gelbart, W. M., & Phillips, R. (2006). Dynamics of DNA Ejection from Bacteriophage. *Biophysical Journal, 91*(2), 411-420. doi:10.1529/biophysj.105.070532

Jiang, L., & Morin, P. J. (2007). Temperature fluctuation facilitates coexistence of competing species in experimental microbial communities. *The Journal of animal ecology, 76*(4), 660-668. doi:10.1111/j.1365-2656.2007.01252.x

Juárez, J. F., Lecube-Azpeitia, B., Brown, S. L., Johnston, C. D., & Church, G. M. (2018). Biosensor libraries harness large classes of binding domains for construction of allosteric transcriptional regulators. *Nature Communications, 9*(1), 3101. doi:10.1038/s41467-018-05525-6

Keen, E. C., Bliskovsky, V. V., Malagon, F., Baker, J. D., Prince, J. S., Klaus, J. S., & Adhya, S. L. (2017). Novel "Superspreader" Bacteriophages Promote Horizontal Gene Transfer by Transformation. *mBio, 8*(1), e02115-02116. doi:10.1128/mBio.02115-16

Kellenberger, E. (2004). The evolution of molecular biology. *EMBO Rep, 5*(6), 546-549. doi:10.1038/sj.embor.7400180

Kerner, A., Park, J., Williams, A., & Lin, X. N. (2012). A programmable Escherichia coli consortium via tunable symbiosis. *PLoS One, 7*(3), e34032. doi:10.1371/journal.pone.0034032

Kim, S., Kwon, S.-H., Kam, T.-I., Panicker, N., Karuppagounder, S. S., Lee, S., . . . Ko, H. S. (2019). Transneuronal Propagation of Pathologic α-Synuclein from the Gut to the Brain Models Parkinson’s Disease. *Neuron, 103*(4), 627-641.e627. doi:10.1016/j.neuron.2019.05.035

Koch, B., Jensen, L. E., & Nybroe, O. (2001). A panel of Tn7-based vectors for insertion of the gfp marker gene or for delivery of cloned DNA into Gram-negative bacteria at a neutral chromosomal site. *J Microbiol Methods, 45*(3), 187-195. doi:10.1016/s0167-7012(01)00246-9

Kotula, J. W., Kerns, S. J., Shaket, L. A., Siraj, L., Collins, J. J., Way, J. C., & Silver, P. A. (2014). Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proc Natl Acad Sci U S A, 111*(13), 4838-4843. doi:10.1073/pnas.1321321111

Ladisch, M. R., & Kohlmann, K. L. (1992). Recombinant human insulin. *Biotechnol Prog, 8*(6), 469-478. doi:10.1021/bp00018a001
Lazonick, W., & Tulum, Ö. (2011). US biopharmaceutical finance and the sustainability of the biotech business model. *Research Policy, 40*(9), 1170-1187. doi:https://doi.org/10.1016/j.respol.2011.05.021
Lim, D., & Song, M. (2019). Development of bacteria as diagnostics and therapeutics by genetic engineering. *Journal of Microbiology, 57*(8), 637-643. doi:10.1007/s12275-019-9105-8
Litchman, E., Edwards, K. F., & Klausmeier, C. A. (2015). Microbial resource utilization traits and trade-offs: implications for community structure, functioning, and biogeochemical impacts at present and in the future. *Front Microbiol, 6*, 254. doi:10.3389/fmicb.2015.00254
Long, S. R. (2001). Genes and Signals in the Rhizobium-Legume Symbiosis. *Plant Physiology, 125*(1), 69. doi:10.1104/pp.125.1.69
Lusetti, S. L., & Cox, M. M. (2002). The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu Rev Biochem, 71*, 71-100. doi:10.1146/annurev.biochem.71.083101.133940
Macarthur, R., & Levins, R. (1967). The Limiting Similarity, Convergence, and Divergence of Coexisting Species. *The American Naturalist, 101*(921), 377-385. Retrieved from www.jstor.org/stable/2459090
Mahr, R., & Frunzke, J. (2016). Transcription factor-based biosensors in biotechnology: current state and future prospects. *Appl Microbiol Biotechnol, 100*(1), 79-90. doi:10.1007/s00253-015-7090-3
Martin, R. G., & Rosner, J. L. (2001). The AraC transcriptional activators. *Current Opinion in Microbiology, 4*(2), 132-137. doi:https://doi.org/10.1016/S1369-5274(00)00178-8
Martínez-García, E., Calles, B., Arévalo-Rodríguez, M., & de Lorenzo, V. (2011). pBAM1: an all-synthetic genetic tool for analysis and construction of complex bacterial phenotypes. *BMC Microbiology, 11*(1), 38. doi:10.1186/1471-2180-11-38
McCall, J. O., Witkin, E. M., Kogoma, T., & Roegner-Maniscalco, V. (1987). Constitutive expression of the SOS response in recA718 mutants of Escherichia coli requires amplification of RecA718 protein. *J Bacteriol, 169*(2), 728-734. doi:10.1128/jb.169.2.728-734.1987
McGinn, J., & Marraffini, L. A. (2019). Molecular mechanisms of CRISPR-Cas spacer acquisition. *Nat Rev Microbiol, 17*(1), 7-12. doi:10.1038/s41579-018-0071-7
McGinn, J., & Marraffini, L. A. (2019). Molecular mechanisms of CRISPR–Cas spacer acquisition. *Nature Reviews Microbiology, 17*(1), 7-12. doi:10.1038/s41579-018-0071-7
McGrady-Steed, J., Harris, P. M., & Morin, P. J. (1997). Biodiversity regulates ecosystem predictability. *Nature, 390*(13), 162-165.
Medicine, I. o. (1995). *Sources of Medical Technology: Universities and Industry*. Washington, DC: The National Academies Press.
Miller, M. B., & Bassler, B. L. (2001). Quorum sensing in bacteria. *Annu Rev Microbiol, 55*, 165-199. doi:10.1146/annurev.micro.55.1.165
Minty, J. J., Singer, M. E., Scholz, S. A., Bae, C.-H., Ahn, J.-H., Foster, C. E., . . . Lin, X. N. (2013). Design and characterization of synthetic fungal-bacterial consortia for direct production of isobutanol from cellulosic biomass. *Proceedings of the National Academy of Sciences, 110*(36), 14592. doi:10.1073/pnas.1218447110
Modell, J. W., Jiang, W., & Marraffini, L. A. (2017). CRISPR-Cas systems exploit viral DNA injection to establish and maintain adaptive immunity. *Nature, 544*(7648), 101-104. doi:10.1038/nature21719
Mukherji, S., & van Oudenaarden, A. (2009). Synthetic biology: understanding biological design from synthetic circuits. *Nature Reviews Genetics, 10*(12), 859-871. doi:10.1038/nrg2697

Munna, M. S., Zeba, Z., & Noor, R. (2016). Influence of temperature on the growth of Pseudomonas putida. *Stamford Journal of Microbiology, 5*(1), 9-12.

Narwani, A., Alexandrou, M. A., Oakley, T. H., Carroll, I. T., & Cardinale, B. J. (2013). Experimental evidence that evolutionary relatedness does not affect the ecological mechanisms of coexistence in freshwater green algae. *Ecology Letters, 16*(11), 1373-1381. doi:10.1111/ele.12182

NIH. How did they make insulin from recombinant DNA? *From DNA to Beer: Harnessing Nature in Medicine & Industry.* Retrieved from https://www.nlm.nih.gov/exhibition/fromdnatobeer/exhibition-interactive/recombinant-DNA/recombinant-dna-technology-alternative.html

Nuñez, J. K., Lee, A. S. Y., Engelman, A., & Doudna, J. A. (2015). Integrase-mediated spacer acquisition during CRISPR–Cas adaptive immunity. *Nature, 519*(7542), 193-198. doi:10.1038/nature14237

Nyholm, S. V., & McFall-Ngai, M. J. (2004). The winnowing: establishing the squid-vibrio symbiosis. *Nat Rev Microbiol, 2*(8), 632-642. doi:10.1038/nrmicro957

Oechslin, F. (2018). Resistance Development to Bacteriophages Occurring during Bacteriophage Therapy. *Viruses, 10*(7), 351. doi:10.3390/v10070351

Olsen, L., Choffnes, E. R., & Mack, A. (2012). *The Social Biology of Microbial Communities*. Paper presented at the Forum on Microbial Threats, Washington, D.C.

Paddon, C. J., & Keasling, J. D. (2014). Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat Rev Microbiol, 12*(5), 355-367. doi:10.1038/nrmicro3240

Papenfort, K., & Bassler, B. L. (2016). Quorum sensing signal-response systems in Gram-negative bacteria. *Nat Rev Microbiol, 14*(9), 576-588. doi:10.1038/nrmicro.2016.89

Philippot, L., Raaijmakers, J. M., Lemanceau, P., & van der Putten, W. H. (2013). Going back to the roots: the microbial ecology of the rhizosphere. *Nature Reviews Microbiology, 11*(11), 789-799. doi:10.1038/nrmicro3109

Ptacnik, R., Solimini, A. G., Andersen, T., Tamminen, T., Brettum, P., Lepistö, L., . . . Rekolainen, S. (2008). Diversity predicts stability and resource use efficiency in natural phytoplankton communities. *Proceedings of the National Academy of Sciences, 105*(13), 5134. doi:10.1073/pnas.0708328105

Purnick, P. E. M., & Weiss, R. (2009). The second wave of synthetic biology: from modules to systems. *Nature Reviews Molecular Cell Biology, 10*(6), 410-422. doi:10.1038/nrm2698

Quianzon, C. C., & Cheikh, I. (2012). History of insulin. *Journal of community hospital internal medicine perspectives, 2*(2), 10.3402/jchimp.v3402i3402.18701. doi:10.3402/jchimp.v2i2.18701

Reis, A., Hornblower, B., Robb, B., & Tzertzinis, G. (2014). CRISPR/Cas9 & Targeted Genome Editing: New Era in Molecular Biology. Retrieved from https://www.neb.com/tools-and-resources/feature-articles/crispr-cas9-and-targeted-genome-editing-a-new-era-in-molecular-biology

Rheinberger, H.-J. (2009). Recent science and its exploration: the case of molecular biology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 40*(1), 6-12. doi:https://doi.org/10.1016/j.shpsc.2008.12.002

Rice, J. J., & Daugherty, P. S. (2008). Directed evolution of a biterminal bacterial display scaffold enhances the display of diverse peptides. *Protein Eng Des Sel, 21*(7), 435-442. doi:10.1093/protein/gzn020

Riglar, D. T., & Silver, P. A. (2018). Engineering bacteria for diagnostic and therapeutic applications. *Nature Reviews Microbiology, 16*(4), 214-225. doi:10.1038/nrmicro.2017.172

Rochat, L., Péchy-Tarr, M., Baehler, E., Maurhofer, M., & Keel, C. (2010). Combination of fluorescent reporters for simultaneous monitoring of root colonization and antifungal gene expression by a biocontrol pseudomonad on cereals with flow cytometry. *Mol Plant Microbe Interact, 23*(7), 949-961. doi:10.1094/mpmi-23-7-0949

Roughgarden, J. (1976). Resource partitioning among competing species—A coevolutionary approach. *Theoretical Population Biology, 9*(3), 388-424. doi:https://doi.org/10.1016/0040-5809(76)90054-X

Sachdeva, G., Garg, A., Godding, D., Way, J. C., & Silver, P. A. (2014). In vivo co-localization of enzymes on RNA scaffolds increases metabolic production in a geometrically dependent manner. *Nucleic Acids Res, 42*(14), 9493-9503. doi:10.1093/nar/gku617

Schleif, R. (2010). AraC protein, regulation of the l-arabinose operon in Escherichia coli, and the light switch mechanism of AraC action. *FEMS Microbiology Reviews, 34*(5), 779-796. doi:10.1111/j.1574-6976.2010.00226.x

Schloss, P. D., & Handelsman, J. (2005). Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. *Genome Biology, 6*(8), 229. doi:10.1186/gb-2005-6-8-229

Scholz, S. A., Diao, R., Wolfe, M. B., Fivenson, E. M., Lin, X. N., & Freddolino, P. L. (2019). High-Resolution Mapping of the Escherichia coli Chromosome Reveals Positions of High and Low Transcription. *Cell Systems, 8*(3), 212-225.e219. doi:https://doi.org/10.1016/j.cels.2019.02.004

Scholz, S. A., Graves, I., Minty, J. J., & Lin, X. N. (2018). Production of cellulosic organic acids via synthetic fungal consortia. *Biotechnol Bioeng, 115*(4), 1096-1100. doi:10.1002/bit.26509

Scott, S. R., Din, M. O., Bittihn, P., Xiong, L., Tsimring, L. S., & Hasty, J. (2017). A stabilized microbial ecosystem of self-limiting bacteria using synthetic quorum-regulated lysis. *Nat Microbiol, 2*, 17083. doi:10.1038/nmicrobiol.2017.83

Sekar, K., Gentile, A. M., Bostick, J. W., & Tyo, K. E. (2016). N-Terminal-Based Targeted, Inducible Protein Degradation in Escherichia coli. *PLoS One, 11*(2), e0149746. doi:10.1371/journal.pone.0149746

Shao, Q., Trinh, J. T., & Zeng, L. (2019). High-resolution studies of lysis–lysogeny decision-making in bacteriophage lambda. *Journal of Biological Chemistry, 294*(10), 3343-3349. doi:10.1074/jbc.TM118.003209

Shapira, P., Kwon, S., & Youtie, J. (2017). Tracking the emergence of synthetic biology. *Scientometrics, 112*(3), 1439-1469. doi:10.1007/s11192-017-2452-5

Sharan, S. K., Thomason, L. C., Kuznetsov, S. G., & Court, D. L. (2009). Recombineering: a homologous recombination-based method of genetic engineering. *Nature protocols, 4*(2), 206-223. doi:10.1038/nprot.2008.227

Sheth, R. U., & Wang, H. H. (2018). DNA-based memory devices for recording cellular events. *Nat Rev Genet, 19*(11), 718-732. doi:10.1038/s41576-018-0052-8

Sheth, R. U., Yim, S. S., Wu, F. L., & Wang, H. H. (2017). Multiplex recording of cellular events over time on CRISPR biological tape. *Science, 358*(6369), 1457-1461. doi:10.1126/science.aao0958

Shong, J., Jimenez Diaz, M. R., & Collins, C. H. (2012). Towards synthetic microbial consortia for bioprocessing. *Curr Opin Biotechnol, 23*(5), 798-802. doi:10.1016/j.copbio.2012.02.001

Silva, K. P. T., & Boedicker, J. Q. (2019). A neural network model predicts community-level signaling states in a diverse microbial community. *PLOS Computational Biology, 15*(6), e1007166. doi:10.1371/journal.pcbi.1007166

Slatkin, M. (1980). Ecological Character Displacement. *Ecology, 61*(1), 163-177. doi:10.2307/1937166

Slouka, C., Kopp, J., Spadiut, O., & Herwig, C. (2019). Perspectives of inclusion bodies for bio-based products: curse or blessing? *Applied microbiology and biotechnology, 103*(3), 1143-1153. doi:10.1007/s00253-018-9569-1

Smith, G. R. (2012). How RecBCD enzyme and Chi promote DNA break repair and recombination: a molecular biologist's view. *Microbiol Mol Biol Rev, 76*(2), 217-228. doi:10.1128/MMBR.05026-11

Song, S., Xu, H., & Fan, C. (2006). Potential diagnostic applications of biosensors: current and future directions. *International journal of nanomedicine, 1*(4), 433-440. doi:10.2147/nano.2006.1.4.433

Starr, M. (2019, 17 August 2019). Scientists Grew a Mysterious Life Form That Could Reveal The Origins of Complex Life. Retrieved from https://www.sciencealert.com/for-the-first-time-scientists-have-grown-the-strange-single-celled-lokiarchaeota-in-a-lab

Stephens, K., Pozo, M., Tsao, C. Y., Hauk, P., & Bentley, W. E. (2019). Bacterial co-culture with cell signaling translator and growth controller modules for autonomously regulated culture composition. *Nat Commun, 10*(1), 4129. doi:10.1038/s41467-019-12027-6

Stewart, E. J. (2012). Growing Unculturable Bacteria. *Journal of Bacteriology, 194*(16), 4151-4160. doi:10.1128/jb.00345-12

Strickland, M. S., Lauber, C., Fierer, N., & Bradford, M. A. (2009). Testing the functional significance of microbial community composition. *Ecology, 90*(2), 441-451. doi:10.1890/08-0296.1

Sutton, T. D. S., & Hill, C. (2019). Gut Bacteriophage: Current Understanding and Challenges. *Frontiers in Endocrinology, 10*(784). doi:10.3389/fendo.2019.00784

Taylor, A. F., & Smith, G. R. (1999). Regulation of homologous recombination: Chi inactivates RecBCD enzyme by disassembly of the three subunits. *Genes & development, 13*(7), 890-900. doi:10.1101/gad.13.7.890

ThermoFisherScientific. Tm Calculator. Retrieved from https://www.thermofisher.com/us/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html

Thomas, B. S., Okamoto, K., Bankowski, M. J., & Seto, T. B. (2013). A Lethal Case of Pseudomonas putida Bacteremia Due to Soft Tissue Infection. *Infect Dis Clin Pract (Baltim Md), 21*(3), 147-213. doi:10.1097/IPC.0b013e318276956b

Thomason, L. C., Costantino, N., & Court, D. L. (2007). *E. coli* genome manipulation by P1 transduction. *Curr Protoc Mol Biol, Chapter 1*, Unit 1 17. doi:10.1002/0471142727.mb0117s79

Thommes, M., Wang, T., Zhao, Q., Paschalidis, I. C., & Segrè, D. (2019). Designing Metabolic Division of Labor in Microbial Communities. *mSystems, 4*(2), e00263-00218. doi:10.1128/mSystems.00263-18

Tilman, D. (1999). THE ECOLOGICAL CONSEQUENCES OF CHANGES IN BIODIVERSITY: A SEARCH FOR GENERAL PRINCIPLES. *Ecology, 80*(5), 1455-1474. doi:10.1890/0012-9658(1999)080[1455:Tecoci]2.0.Co;2

Toms, A., & Barrangou, R. (2017). On the global CRISPR array behavior in class I systems. *Biology direct, 12*(1), 20-20. doi:10.1186/s13062-017-0193-2

Toyofuku, M., Nomura, N., & Eberl, L. (2019). Types and origins of bacterial membrane vesicles. *Nat Rev Microbiol, 17*(1), 13-24. doi:10.1038/s41579-018-0112-2

Tran, F., & Boedicker, J. Q. (2019). Plasmid Characteristics Modulate the Propensity of Gene Exchange in Bacterial Vesicles. *Journal of Bacteriology, 201*(7), e00430-00418. doi:10.1128/JB.00430-18

Valero-Rello, A., López-Sanz, M., Quevedo-Olmos, A., Sorokin, A., & Ayora, S. (2017). Molecular Mechanisms That Contribute to Horizontal Transfer of Plasmids by the Bacteriophage SPP1. *Frontiers in Microbiology, 8*, 1816-1816. doi:10.3389/fmicb.2017.01816

van Elsas, J. D., Chiurazzi, M., Mallon, C. A., Elhottova, D., Kristufek, V., & Salles, J. F. (2012). Microbial diversity determines the invasion of soil by a bacterial pathogen. *Proc Natl Acad Sci U S A, 109*(4), 1159-1164. doi:10.1073/pnas.1109326109

Vandermeer, J. H. (1975). Interspecific competition: a new approach to the classical theory. *Science, 188*(4185), 253-255. Retrieved from https://science.sciencemag.org/content/sci/188/4185/253.full.pdf

Vigneshvar, S., Sudhakumari, C. C., Senthilkumaran, B., & Prakash, H. (2016). Recent Advances in Biosensor Technology for Potential Applications – An Overview. *Frontiers in Bioengineering and Biotechnology, 4*(11). doi:10.3389/fbioe.2016.00011

Volterra, V. (1928). Variations and Fluctuations of the Number of Individuals in Animal Species living together. *ICES Journal of Marine Science, 3*(1), 3-51. doi:10.1093/icesjms/3.1.3

von Bodman, S. B., Willey, J. M., & Diggle, S. P. (2008). Cell-Cell Communication in Bacteria: United We Stand. *Journal of Bacteriology, 190*(13), 4377. doi:10.1128/JB.00486-08

von Wintersdorff, C. J. H., Penders, J., van Niekerk, J. M., Mills, N. D., Majumder, S., van Alphen, L. B., . . . Wolffs, P. F. G. (2016). Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. *Frontiers in Microbiology, 7*(173). doi:10.3389/fmicb.2016.00173

Watson, J. D., & Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature, 171*(4356), 737-738. doi:10.1038/171737a0

Weathers, P. J., Elkholy, S., & Wobbe, K. K. (2006). Artemisinin: The biosynthetic pathway and its regulation in Artemisia annua, a terpenoid-rich species. *In Vitro Cellular & Developmental Biology - Plant, 42*(4), 309-317. doi:10.1079/IVP2006782

Wigley, D. B. (2013). Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB. *Nat Rev Microbiol, 11*(1), 9-13. doi:10.1038/nrmicro2917

Yehl, K., Lemire, S., Yang, A. C., Ando, H., Mimee, M., Torres, M. T., . . . Lu, T. K. (2019). Engineering Phage Host-Range and Suppressing Bacterial Resistance through Phage Tail Fiber Mutagenesis. *Cell, 179*(2), 459-469 e459. doi:10.1016/j.cell.2019.09.015

You, L., Cox, R. S., 3rd, Weiss, R., & Arnold, F. H. (2004). Programmed population control by cell-cell communication and regulated killing. *Nature, 428*(6985), 868-871. doi:10.1038/nature02491

Yurtsev, E. A., Conwill, A., & Gore, J. (2016). Oscillatory dynamics in a bacterial cross-protection mutualism. *Proc Natl Acad Sci U S A, 113*(22), 6236-6241. doi:10.1073/pnas.1523317113

Zhang, H., Pereira, B., Li, Z., & Stephanopoulos, G. (2015). Engineering Escherichia coli coculture systems for the production of biochemical products. *Proc Natl Acad Sci U S A, 112*(27), 8266-8271. doi:10.1073/pnas.1506781112

Zhang, S., Merino, N., Okamoto, A., & Gedalanga, P. (2018). Interkingdom microbial consortia mechanisms to guide biotechnological applications. *Microb Biotechnol, 11*(5), 833-847. doi:10.1111/1751-7915.13300

Zhou, K., Qiao, K., Edgar, S., & Stephanopoulos, G. (2015). Distributing a metabolic pathway among a microbial consortium enhances production of natural products. *Nat Biotechnol, 33*(4), 377-383. doi:10.1038/nbt.3095

Zhou, S., Gravekamp, C., Bermudes, D., & Liu, K. (2018). Tumour-targeting bacteria engineered to fight cancer. *Nature reviews. Cancer, 18*(12), 727-743. doi:10.1038/s41568-018-0070-z

Ziv, N., Brandt, N. J., & Gresham, D. (2013). The use of chemostats in microbial systems biology. *J Vis Exp*(80). doi:10.3791/50168