Genetic Technologies to Engineer and Understand the Microbiome

by

Mark Mimee

BSc. in Microbiology & Immunology, McGill University (2011)

Submitted to the Microbiology Graduate Program
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2018

Signature redacted

Author . . . .
Microbiology Graduate Program

Signature redacted May 25, 2018

Certified by . . . .
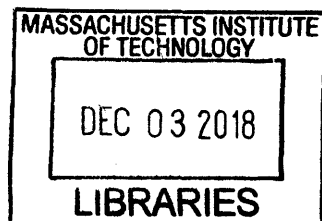Timothy K. Lu
Associate Professor of Biological Engineering and Electrical
Engineering and Computer Science
Thesis Supervisor

Signature redacted

Accepted by . . . .
Kristala L. Jones Prather
Arthur D. Little Professor of Chemical Engineering
Chair of Microbiology Program

Genetic Technologies to Engineer and Understand the Microbiome

by

Mark Mimee

Submitted to the Microbiology Graduate Program
on May 25, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The microbes that inhabit the human body are integral to human health and disease: from inflammatory bowel disease to allergy, metabolic syndrome to autism. Due to its high connectivity with human physiology, manipulation of the microbiota has therapeutic potential in a vast array of diseases. However, techniques for targeted modification of microbial communities are currently lacking. In this thesis, I present several technologies that can be applied to engineer and better understand the microbiota. First, we present a subtractive strategy for microbiota manipulation using CRIPSR-Cas engineered bacteriophage that can selectively remove target strains from a community based on the presence of target DNA sequences. Next, we describe an additive strategy whereby commensal *Bacteroides* spp. are genetically modified to perform novel functions within the murine microbiota. We developed a suite of genetic parts to facilitate organism design and engineering. These tools were then expanded to engineer outer membrane vesicles derived from *Bacteroides* as immunomodulatory agents. Finally, we leveraged the natural sensing abilities of bacteria to create cellular biosensors for biomarkers of gastrointestinal disease. Heme biosensors were paired with readout electronics to generate an ingestible medical device for *in situ* detection of gastrointestinal bleeding. The technologies described herein contribute to the progression of microbiome engineering towards clinical applications and the advancement of our understanding of how our smallest friends impact our health.

Thesis Supervisor: Timothy K. Lu
Title: Associate Professor of Biological Engineering and Electrical Engineering and Computer Science

Acknowledgements

One lesson that has rung true throughout my tenure at MIT is that science is best done in teams. Although there is only one name on the cover of this thesis, this work represents the aggregate contribution of the wide range of support I have received throughout my PhD. From sage mentorship to long afternoons of watching mice defecate, from jubilant celebrations to brutal failure after failure after failure, from vibrant brainstorming sessions to repeated transfer liquid from one tube to another, I have been accompanied by mentors, colleagues, friends, and family throughout my journey through graduate school. I am sincerely grateful for the relationships that have made this journey possible.

Firstly, I would like to thank Timothy Lu for his support and guidance in his laboratory. As an early member of the Lu lab, I've had the privilege to watch the lab rocketship from a small off-brand, biology lab in MIT Electrical Engineering to a premier synthetic biology lab. My time in Tim's lab has opened my eyes to multiple sides of science: basic research, communications, entrepreneurship, among others. I thank Tim for affording me the independence to drive my own research, for allowing me the space to fail, and to learn from those failures. I am grateful for the infectious excitement he injects into projects and the motivation to continue achieving. Finally, I appreciate the confidence he has inspired in me to pursue opportunities outside of my comfort zone and to continue to try to learn as much as possible.

My collaborating advisor Anantha Chandrakasan offered me a walk through the world of electrical engineering. I thank him for the opportunity to learn from his astute direction and keen expertise. My collaborative work with Christopher Voigt gave me a valued perspective of synthetic biology. I value his ability to craft a project

5

Jasper van der Peet, and Djenet Bousbaine, inspired me to persevere on a challenging project despite abundant failures. I'm fortunate to have worked with researchers and clinicians that shared in a common vision. I especially *appreciate* Logan's efforts as a friend, a co-worker, and a sober clinical voice in a room of basic researchers. My biosensor work was conceived hand-in-hand with fellow proud Canadian, Phillip Nadeau. An impressive semiconductor chip designer with a refreshing love of learning, Phil introduced me to the exciting world of electronics and taught me the value of cross-disciplinary work. Our journeys to San Diego, Japan, and Hong Kong are some of my most memorable memories from MIT. Finally, I would also like to express my gratitude working alongside Robert Citorik, both with CRISPR-phage and with continued discussions throughout graduate school. A colleague in Micro 2011, we endured all of the ups and downs of graduate school and mid-late twenties life together. I am truly thankful to have such a close friend throughout my stay at MIT.

My family's support has been unwavering during the completion of this thesis and I would like to express my sincere thanks. To Grandma, a shining example of a true carer. To my big sister, Andrea, my inspiration to pursue a career in science and an example of confidence, tenacity, and humour. To my father, who taught me patience, to think outside of the box, to keep my head up and my stick on the ice, and to find laughter in all aspects of life. To my mother, who instilled in me the values of hard work and unconditional caring, fostered a deep love of learning, and showed me the importance of communication and constructive debates through our epic English essay arguments. The lessons I have learned from my family are innumerable and form the intellectual and philosophical framework that propelled the work described herein.

With me for the entirety of my PhD, the continual love and support of my wife, CJ, has kept my spirits and motivation high. I am profoundly grateful to have such a supportive spouse able to tolerate my consistent tardiness, all-nighters pulled in a small studio apartment, and incoherent ramblings about my experiments. I call myself fortunate to receive encouragement for personal growth as well as a safe environment to take risks. Thank you for being my key uncredited collaborator on this work and

I look forward to future projects.

Contents

List of Figures

Chapter 1

Introduction

The human body is home to a diverse community of symbiotic, commensal and pathogenic microorganisms, collectively known as the microbiota. These bacterial, viral and eukaryotic communities exist on all environmentally exposed sites in the body, including the skin, nasopharynx, oral cavity, respiratory tract, gastrointestinal tract and female reproductive tract. Community composition is largely phylogenetically conserved across healthy individuals for a given body site [1]. The intimate association between host and microbe over the course of a lifetime has profound implications on human health, across multiple physiological axes. Indeed, these natural beneficial interactions make the microbiota an attractive avenue for engineering human physiology. In the following chapter, I will summarize illustrative host-commensal interactions linked with various physiological processes. Additionally, I will outline the synthetic biology efforts that have sought to manipulate the microbiota for diagnostic or therapeutics purposes. This review should provide a conceptual foundation of the importance of the microbiota on human health and the state of current technologies that take advantage of host-commensal interactions. Finally, I will outline the scope and organization of the following chapters.

Figures and text in the following introduction have been adapted from the 2016 Advanced Drug Delivery review entitled 'Microbiome therapeutics – Advances and Challenges' to which I contributed as a first author [2].

1.1 Effects of the Gut Microbiota on the Host

A major focus of human microbiome research has been studying the bacteria in the gut, which represent the largest community both in terms of abundance and diversity [1]. Initial colonization occurs at birth and the mode of delivery (i.e. vaginal or Caesarian section) influences the founding community [3]. Early-life events, such as transitions in diet [4] and antibiotic use [5], shape the volatile infant microbiome, which stabilizes with age. In contrast, the adult gut microbiota is highly resilient to minor perturbations. For instance, the individual gut microbiome was determined to be 60% conserved at the strain level over the course of five years, with members of the phyla Bacteroidetes and Actinobacteria being the most stable [6]. This longitudinal stability, in combination with the great interpersonal diversity of the microbiome, even permits identification of >80% of individuals by their unique 'microbial fingerprint' [7]. Despite its stability, large insults, such as antibiotic treatment [8] or the onset of disease [9], can lead to dramatic changes in the composition of the gut flora. Additionally, diet and environment are major contributing factors shaping the community [10, 11]. Studies in humans [12] and mice [13–15] have shown that the microbiota responds to major changes in diet that reflect nutritional intake. In a large Flemish and Dutch study, other environmental factors such as body-mass index, stool disease, disease, pharmaceutical intake and age have significant effects on the composition of the microbiota [16]. An Israeli study echoed many of these factors and added HDL cholesterol levels, lactose intake and glycemic status to the list [17]. Of note, with increased statistical power, the same Israeli study demonstrated the minimal effect human genetics and specific single-nucleotide polymorphisms have on the human microbiota. Despite these large studies consisting of thousands of human samples across geographically distinct populations, the measured environmental factors only account for 20% of the variance between individuals microbiota pointing to a suite of unknown or stochastic factors that exert a profound influence [16, 17].

Small-molecules produced by members of the microbiota have drawn considerable attention from both the microbiome and chemical biology communities as ma-

jor players in host-microbe interaction. Traditionally, small molecules derived from soil microorganisms have been a mainstay of pharmaceutical drug discovery and the source of several blockbuster antimicrobial and anticancer agents [18]. With millennia of co-evolution with mammalian hosts, it is expected that microbiota-derived natural products will similarly host a treasure trove of bioactive molecules. Human metagenomes are replete with biosynthetic gene clusters that can be identified computationally [19]. Mining these clusters for bioactivity has revealed molecules that act as antimicrobials [19], G-protein-coupled receptor ligands [20] , and protease inhibitors [21]. Advancements in DNA synthesis and synthetic biology has enabled the de novo synthesis of these biosynthetic gene clusters [22] and characterization of the synthesized molecules. Identification of the molecular players in host-commensal interaction will surely be a powerful tool in capitalizing on the microbiota for therapeutic applications.

1.1.1 Metabolism

At the community level, the collection of microbes that inhabit the gut plays a key role in human metabolism. Early observations recognized the relative proportions of Bacteroidetes to Firmicutes as a microbial signature of obesity in mice, where an obese-associated microbiota displayed an increased capacity for energy harvest [23, 24]. Notably the increase in total body and fat mass, as well as the associated metabolic phenotypes, was transferrable by transplantation of obese-associated flora into gnotobiotic mice [24, 25]. Similarly, studies on the gut microbiota of twins discordant for kwashiorkor, a severe form of acute malnutrition, demonstrated that transplantation of a kwashiorkor-associated microbiota into gnotobiotic mice leads to a rapid decrease in body weight and metabolic phenotypes associated with malnutrition [26]. Later studies have linked specific microbial taxa with lean or obese phenotypes. For example, murine *Bacteroides acidifaciens* [27] as well as *Akkermansia muciniphila* [28] can protect against obesity and insulin resistance in mice fed high fat diets. Conversely, branched amino acids produced by *Prevotella copri* and *Bacteroides vulgatus* have been linked with insulin resistance in both murine and human

studies [29].

At the molecular level, several microbially-derived metabolites have been shown to directly or indirectly modulate host metabolism. A major metabolic role of the gut flora is the conversion of ingested dietary fiber and host mucosal glycans into short-chain fatty acids (SCFAs), which include acetate, propionate and butyrate [30]. SCFAs mediate a multitude of beneficial effects on the host. Butyrate serves as a primary energy source to colonic epithelial cells [31, 32] and increases gut barrier integrity through stabilization of hypoxia-inducible transcription factor [33]. Moreover, SCFAs can modulate host gene expression through inhibition of histone deacetylases [34] or through G-protein-coupled cell surface receptor signaling [35, 36]. This signaling stimulates intestinal gluconeogenesis [37], secretion of glucagon-like peptide 1, leptin, and peptide YY [36, 38], and modulates gut motility [36, 39]. In addition to SCFAs, microbial metabolism of bile salts can also influence gut motility [14]. Furthermore, the gut microbiota, in concert with hepatic enzymatic activity, participates in the conversion of dietary phosphatidylcholine and L-carnitine into trimethylamine N-oxide (TMAO), a molecule strongly correlated with atherosclerotic heart disease [40, 41]. TMAO production is dependent on an intact microbiota [41] and transferrable by microbiota transplantation [42]. Inhibition of microbial enzymes involved in TMAO production has been shown to be an effective means of treating atherosclerosis in mice, suggesting that inhibitors of microbial enzymatic activity could be potential therapeutics for host diseases [43].

Just as the gut microbiota metabolizes carbohydrates, proteins and lipids ingested by the host, bacteria in the gut interact with compounds foreign to living organisms, known as xenobiotics [8]. While these interactions often only result in mild changes in gene expression in gut microbes [8], a high-throughput survey of human-targeted drugs effects on human commensal species revealed a extensive in vitro growth inhibition [44]. These direct effects of human-targeted drugs on bacterial growth was concordant clinical data of metformin [45] and protein-pump inhibitor [46, 47] modulation of microbiome composition . In contrast, microbial activity can disrupt the intended activity of host-targeted drugs. For example, *Eggerthella lenta*, a gut Acti-

nobacteria, can inactivate the cardiac drug digoxin [48] and repression of the responsible operon through supplementation of dietary protein can increase serum levels of active digoxin [49]. Additionally, bacterial β-glucuronidases interfere with the efficacy of the colon cancer chemotherapeutic CPT-11 by reactivating the drug in the intestine, resulting in an increased toxicity that can be reversed through small molecule inhibition of β-glucuronidase activity [50]. Together, the repercussions of microbial metabolic activity on both nutrients and xenobiotics impact human health.

1.1.2 Immunity

The homeostatic association of the human body with trillions of microbes demands a continuous dialogue between the host immune system and the microbial flora [51–53]. In both the small and large intestine, commensals are kept in check sthrough a thick mucosal layer, gradients of antimicrobial peptides and mucosal IgA antibodies [51, 54]. Nonetheless, the gut flora and their associated metabolites strongly influence the immune system in both health and disease [55, 56]. Inflammatory bowel disease (IBD) is thought to be caused by a combination of environmental, genetic and microbial factors [54]. Indeed, there is a strong association between constellations of microbial taxa in patients with IBD compared to healthy subjects [9, 57]. Certain microbially derived products, such as polysaccharide A from *Bacteroides fragilis* [58, 59] and butyrate from *Clostridia* spp. [60–62], have been shown to be protective in colitis through the induction and maintenance of peripheral regulatory T cells (T_{REG}). Similarly, protein antigens produced by segmented filamentous bacteria have been demonstrated to robustly induce T_{H17} responses [63] [64]. While these microbe-host immune interactions are predominant in the gut, the effects can manifest systemically. For example, a depletion in certain microbial taxa in the infant gut, which include *Lachnospira*, *Veillonella*, *Faecalibacterium*, and *Rothia*, is a risk factor for childhood asthma [65], while colonization of mice with spore-forming *Clostridia* species can protect against food allergy sensitization [66]. Early colonization with poorly immunogenic bacteria, such as *Bacteroides* spp., can also disrupt the proper development of the immune system and lead to autoimmune disease [56]. The inter-

action between microbe and host is a powerful determinant in the development of a healthy, balanced immune system.

Bacteria in the gut can also indirectly modulate immune function through interactions with invading pathogens [67, 68]. The capacity of the normal flora to defend against pathogenic invasion is referred to as colonization resistance and is founded on either competition for limited resources or microbial warfare. Ablation of the gut microbiota through broad-spectrum antibiotics predisposes individuals to opportunistic pathogens, including *Clostridium difficile*, that are normally kept at bay by commensal microbes [69]. Endogenous and probiotic strains of *Escherichia coli* can protect against *Salmonella* Typhimurium infection in mice by scavenging iron necessary for a productive infection [70]. Furthermore, the presence of *Clostridium scindens* can defend against recurrent *C. difficile* infections through the synthesis of secondary bile acids toxic to *C. difficile* [71]. Bacteria also possess an arsenal of molecular weapons, such as bacteriocins, microcins and Type VI secretion systems [72], which exclude competing invaders from a community [68]. While antagonistic interactions help prevent pathogens from entering the microbiota, some interactions have been shown to exacerbate infection. For example, in mono-colonized mice, *Bacteroides thetaiotaomicron* can forage sialic acid from mucosal glycans and generate succinate, which can then serve as a carbon source for *C. difficile* during infection [73, 74]. Thus, bacteria-bacteria interactions can have consequences for the host in which they occur.

1.1.3 Gut-Brain Axis

Strikingly, the impact of the gut flora on the host reaches beyond the gastrointestinal tract to the brain, where microbially derived metabolites can influence neuroendocrine function and behaviour [75–77]. As mentioned above, butyrate produced by the microbiota modulates levels of neuropeptides that govern satiety, such as leptin and peptide YY [36]. Additionally, the presence of a microbiota as well as microbiota-produced SCFAs worsen pathology in murine Parkinson's models [78]. The metabolic activity of bacteria can also influence circulating levels of neurotransmitters, such as serotonin [79], and bacterial enzymes are capable of directly synthesizing the neuro-

transmitter tryptamine [80]. Together, these endocrine effects can manifest in changes in behaviour. In mouse models of autism spectrum disorders, correction of a dysbiotic microbiota through administration of *B. fragilis* can lead to a reversal in behavioural abnormalities [81].

1.2 Efforts to Harness and Engineer the Microbiota
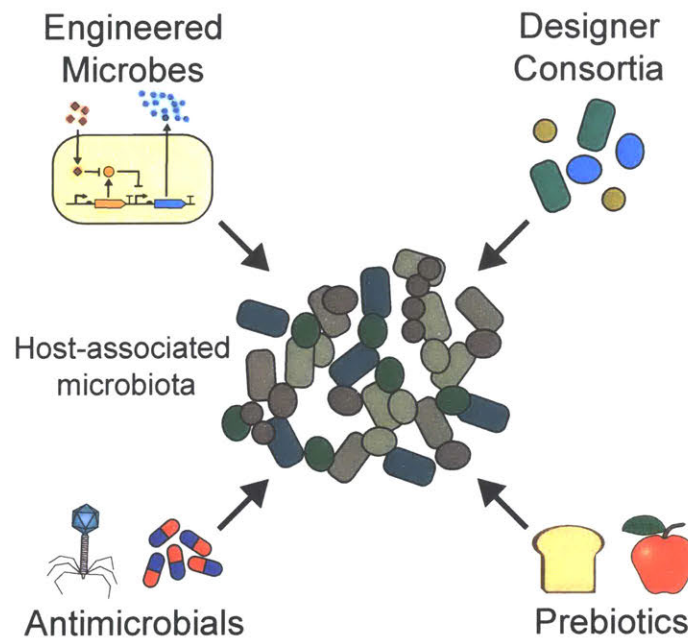
Figure 1-1: **Approaches to Microbiome-Based Therapeutics**: Additive therapies involved supplementation of the microbiota with one or more natural or engineered microbes. Subtractive therapies employ antimicrobial agents to selectively remove target strains from the community. Modulatory therapies use abiotic agents to alter the composition or activity of endogenous microbiota strains.

As a decade of research has solidified the importance of the microbiome on human health and disease, recent efforts have sought to capitalize on this interaction to create microbiome-based therapeutics. These approaches can be classified into three major paradigms: additive, subtractive or modulatory therapies (Fig. 1-1). Additive

therapy includes the supplementation of the host microbiota with either individual strains (Fig. 1-2) or a consortium (Fig. 1-3) of natural or engineered microorganisms. Subtractive therapy refers to the specific elimination of deleterious members of the microbiome to cure disease (Fig. 1-4). Finally, modulatory therapies involve administration of non-living agents, or prebiotics, to modulate the composition or activity of the endogenous microbiome. Probiotics and prebiotics constitute the first generation of microbiome therapies and have been reviewed extensively [82–86]. Here, I will focus on the next-generation of microbiome therapeutics using recombinant probiotics, designed microbial consortia and selective antimicrobials. Thus far, therapeutic bacterial consortia have been primarily composed of natural strains, whereas genetically engineered bacteria have mainly been delivered as individual strains. However, moving forward, genetically modified communities may be developed that can incorporate the diversity and robustness of microbial consortia with the added efficacy and controllability of synthetic gene circuits.

1.2.1 Genetically Engineered Probiotics
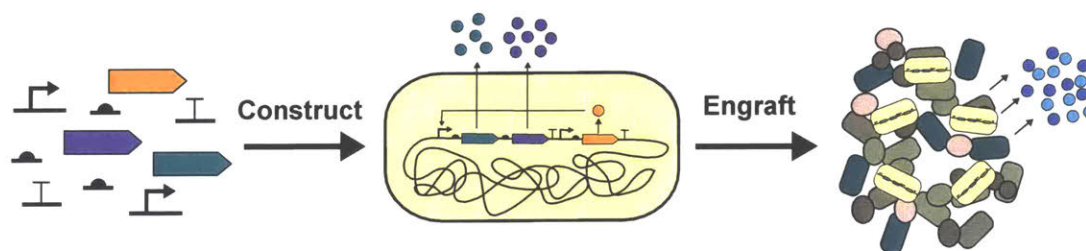
Figure 1-2: **Genetically Engineered Probiotics as Therapeutic Agents**: Gene circuits are constructed using libraries of genetic parts to enable microbial production of therapeutic proteins. Introduction of these microorganisms into the endogenous microbiota allows for *in situ* detection of disease biomarkers and/or drug production.

Probiotic therapies are founded on the rationale that naturally occurring human-associated microbes provide a myriad of health benefits. Indeed, oral ingestion of *Lactobacillus* spp., *E. coli* and *Bifidobacterium* spp. has been clinically shown to remedy a wide variety of diseases [87–90]. Recent efforts seek to augment the natural benefits of probiotics through recombinant expression of therapeutic biomolecules.

Implementation of cells as vehicles of drug delivery could enable continuous *in situ* production of biotherapeutics that could overcome issues such as bioavailability and drug inactivation associated with oral administration. Synthesis of protein therapeutics could be conditioned upon the detection and integration of specific environmental cues related to disease. This conditional, on-demand drug release is a particularly attractive advantage of cell-based therapeutics that could enable entirely new pharmacological paradigms. If the therapeutic organism can stably colonize the host, the engineered microbe could dynamically correct perturbations caused by disease to restore homeostasis in the host. Realization of this vision of fully autonomous 'smart' cell-based therapeutics is still forthcoming. However, recent efforts have demonstrated the efficacy of cell-based bacterial therapeutics in preventing infection, resolving inflammation and treating metabolic disorders.

Natural colonization resistance conferred by native members of the normal flora has been augmented through cellular engineering. Duan et al. explored the prophylactic use of probiotic *E. coli* Nissle 1917 engineered to inhibit virulence of *Vibrio cholerae* in infant mouse models [91]. Virulence of *V. cholerae* is coordinated by extracellular quorum sensing molecules, which modulate gene expression in a density-dependent manner. When bacterial numbers are low, *V. cholerae* expresses virulence factors necessary for the establishment of infection; when numbers are high, virulence factor expression is repressed to allow escape from the host. *E. coli* was engineered to crosstalk with quorum sensing and thereby inhibit productive infection. Administration of therapeutic cells resulted in an increase in survival and a concomitant decrease in bacterial burden and cholera toxin expression. Similarly, Hwang et. al engineered *E. coli* Nissle 1917 to sense the presence of *Pseudomonas aeruginosa* and secrete antimicrobial proteins to combat the pathogen [92]. In mouse models of decolonization, the engineered strains of *E. coli* could reduce the *P. aeruginosa* burden by several orders of magnitude. Additionally, in the context of the genital tract, *Lactobacillus jensenii* was genetically modified to prevent transmission of chimeric simian/human immunodeficiency virus (SHIV) in a rhesus macaque model [93]. Bacteria were engineered to produce the antiviral cyanovirin-N. Prophylactic treatment of macaques

decreased both the occurrence of SHIV following multiple challenges as well as the peak viral load. Engineered bacteria could thus serve as treatments for both bacterial and viral infections.

IBD has drawn particular attention as a compelling candidate for cell-based therapeutics, due to the implications of the gut microbiota in disease and the lack of long-term, cost-effective treatments [54]. Early work explored the use of *Lactococcus lactis* engineered to constitutively secrete recombinant interleukin-10 (IL-10), a potent anti-inflammatory cytokine depleted in IBD [94]. Using both chemically and genetically induced models of colitis, Steidler et al. demonstrated that recombinant microorganisms could be implemented to reduce pathology and suppress pro-inflammatory cytokine secretion. Phase I clinical trials also showed that recombinant *L. lactis* treatment was well-tolerated in a small Crohn's disease cohort, although efficacy in treating symptoms was modest [95]. IL-10-secreting *L. lactis* were further modified to produce either the auto-antigen proinsulin [96] or glutamic acid decarboxylase-65 [97] for treatment of autoimmune diabetes. When administered in conjunction with anti-CD3 therapy, both recombinant organisms were able to induce tolerance, increase numbers of regulatory T cells and reverse hyperglycemia in mice [96, 97]. Microbial expression of other anti-inflammatory cytokines, such as transforming growth factor-β1 [98] and anti-tumor necrosis factor α Nanobodies [99], as well as the tissue repair factor keratinocyte growth factor-2 [100], have been shown to be protective against colitis in mouse models of IBD. In addition, production of the protease inhibitor Elafin from lactic acid bacteria was shown to restore proteolytic homeostasis in mouse models of colitis and protect against inflammation [101]. Recombinant bacteria have also been implemented for treating oral mucositis, a condition involving ulcerative lesions that is a common side effect of chemotherapy. Topical application of *L. lactis* engineered to secrete trefoil factor-1 was effective at treating oral mucositis in hamster models [102] and early clinical trial data indicate that treatment is well-tolerated and could be effective at reducing incidence [103]. Together, these studies suggest that recombinant cellular therapies could be viable therapeutic agents to treat inflammatory diseases.

Integration of recombinant microbes into the host microbiome has also seen success in treating metabolic diseases, such as obesity and diabetes. Daily feeding of probiotic *E. coli* engineered to synthesize precursors of anorexigenic lipids reduced obesity, adiposity and food intake in mice fed a high-fat diet [104]. These protective effects were sustained weeks after cessation of bacterial treatment [104]. Moreover, *Lactobacillus gasseri* was used as a vehicle for delivery of GLP-1, a protein able to induce conversion of intestinal epithelial cells into insulin producing cells [105]. Administration of the engineered probiotic increased the numbers of intestinal insulin-producing cells and decreased hyperglycemia in a rat model [105].

The microbiota could also serve as a useful tool for diagnosis of disease. Indeed, as the composition of the microbiota is modulated in specific manners in the context of disease, 16S sequencing profiles can provide insight into the presence or absence of disease. For example, microbiome sequencing data, in concert with fecal immunochemical tests, can be helpful in the diagnosis of colonic lesions and adenomas [106]. Response to biologic therapy for Crohn's disease patients can also be measured with fecal microbiota signatures [107]. Just as engineering efforts have engendered several proof-of-concept examples of therapeutic bacteria, synthetic biology groups have developed microbes that can act as diagnostics of human disease. In these systems, small molecule-responsive transcription factors are co-opted to drive expression of reporter genes. More sophisticated gene circuits can even be employed for signal amplification or recording of exposure to signal [108]. Engineered microbes hold great promise for *in situ* detection of labile biomarkers that may degrade before sampling can occur. In early efforts, Kotula et. al [109] engineered an *E. coli* strain to record exposure to a model inducer anhydrotetracycline (aTc) in the mouse gastrointestinal tract. More recent efforts have employed similar technologies to record exposure to tetrathionate or thiosulfate, sulfur metabolites generated during oxidative stress in the context of colitis [110, 111]. While biosensor-based diagnostics are still in their infancy, the potential to combine the sensing capabilities of microorganisms with the capacity for therapeutics production could yield smart, adaptive cell-based therapeutics.

1.2.2 Engineered Consortia
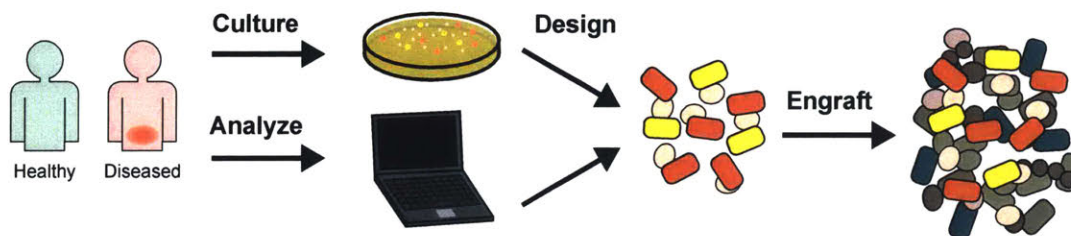
Figure 1-3: **Engineered Consortia for Additive Microbiome Therapies**: Designer microbial consortia are informed by community profiling studies of clinical samples from healthy and diseased patients. Clinical isolates from these patients can then be assembled into a defined mixture of microorganisms that can reprogram the microbial ecology within an individual.

Engineering the microbial community as a whole is an alternative strategy that has seen tremendous clinical success for treatment of recurrent *C. difficile* infections. Infusion of stool from healthy donors to diseased patients, termed a fecal microbiota transplant, boasts greater than 90% efficacy in resolving recurrent infections [112] and is more than twice as effective as antibiotic treatment alone [113]. Despite their clinical success, fecal microbiota transplants draw safety concerns for fear of introducing pathogens or opportunists that could exacerbate disease [114]. A regulatory framework and strict screening guidelines for donors has been established [114], but deciphering the minimal subset of therapeutic microbes has been a focus for mitigating safety concerns and increasing the reliability of treatments [115]. Aside from recurrent *C. difficile* infections, many believe fecal microbiota transplants hold promise in treating IBD [116] and early trials showed modest success [117, 118]. However, due to the more complex nature of disease and a higher incidence of adverse events, more trials are necessary to establish stool transplants or infusions of defined microbial communities as a viable treatment option for IBD [117].

In mouse models, microbiota reconstitution has proved successful in altering community-wide metabolic activity of urea [119]. In the gut, urea produced by the liver is converted to ammonia and carbon dioxide by bacterial ureases. Accumulation of systemic ammonia is associated with neurotoxicity and encephalopathy in patients with liver

deficiency. Following depletion of the endogenous microbiota using antibiotics and polyethylene glycol as a purgative, Shen et al. could transplant a defined microbial community with low urease activity that remained stable for months [119]. In a hepatic injury model, the redefined microbiota increased survival and protected against cognitive defects associated with hyperammonemia. This study demonstrates the feasibility in rationally sculpting a host-associated microbial community to protect against metabolic diseases.

1.2.3   Subtractive Approaches
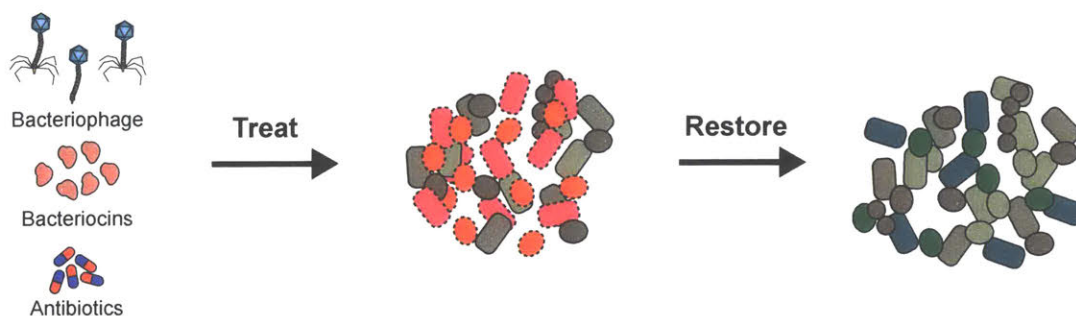
Figure 1-4: **Subtractive Approaches in Microbiome Therapeutics**: Bacteriophages, bacteriocins and small molecule antibiotics can be used to selectively eliminate deleterious microbes from the microbiota. Consequently, the loss of specific taxa elicits a global shift in the microbial community as new constituents occupy the niches of the eliminated microbes. The addition of engineered bacteria together with selective elimination of targeted strains may provide enhanced therapeutic efficacy.

Subtractive therapies may employ chemicals, peptides, or even replicating entities to remove bacteria from the gut with varying degrees of specificity. In medicine, this is currently accomplished through the use of antibiotics, which tend to be broad-spectrum in nature, exhibiting activity towards many different bacteria. As a result, treatment of a patient aimed to remove an infectious pathogen also leads to the unintended reduction of other members of the microbiota. This community shifting may cause the patient to become susceptible to other temporary or chronic conditions to which they are normally protected, including antibiotic-associated infections with *C. difficile*. The development of targeted antimicrobials, such as bacteriocins and

bacteriophage, could yield more effective subtractive therapies.

Bacteria that occupy overlapping niches need to compete for the same resources. Bacteriocins are ribosomally synthesized antimicrobial peptides that are produced by bacteria and may help gain an edge over competitors by exhibiting toxic activity towards susceptible cells. A producer cell encodes both toxic and immunity functions, killing neighbouring cells while protecting itself and its progeny from the effects of the bacteriocin. Recent studies have used metagenomics to identify bacteriocins [120] [121], shown potential fitness advantages of producers *in vivo* [122], and observed protective effects of producers against pathogens [123, 124]. As these molecules seem to play a role in modulating microbial populations within a host, they may prove useful tools for subtractive-type therapies.

Bacteriophage therapy is a highly specific method of killing bacteria through the use of natural or engineered viral parasites. Bacteriophages, or phages, are viruses that infect bacteria and use cellular resources to produce progeny, generally killing the host cell in the process. Discovered approximately 100 years ago [125, 126], the application of phages as antimicrobials has seen a renewed interest with the growing threat of antibiotic-resistant pathogens [127, 128]. Though early phage therapies targeted intestinal pathogens [129], clearance issues have recently been reported wherein bacterial and phage populations stably coexist in the murine gut [130, 131]. Knowledge attained from research into the ecology of phages in the gut may be pivotal in determining factors that lead to successful therapies in this complex ecosystem. A newer focus of the field is to examine the natural role of these viruses in shaping host-associated bacterial populations [132]. Metagenomic research of the human microbiome has described the fecal virome of both healthy and diseased donors. These studies include measuring phage diversity, variability, and stability [133], and analyzing changes associated with diet [134] or IBD [135] in humans, or antibiotic treatment in mice [136] . These types of studies should yield important information to be used for designing phage-based strategies for targeting pathogens or altering microbial communities in the gut.

Engineered phages, or those modified to carry additional or alternative functions

to those naturally occurring, may prove useful for therapeutics as design can be informed by new knowledge. Recently, it was found that certain phages possess Ig-like protein domains on their capsids that enhance association with mucus [137]. This or alternative localization domains may be useful for improving residence time in the gut or helping concentrate phages to relevant biogeographies. Phage engineering efforts have included altering host adsorption factors to change phage host range [138] or encoding a dispersal enzyme to help break up bacteria in protective biofilms [139]. Phage have also been used as DNA delivery agents to reverse antibiotic resistance [140, 141] or to exert broad-spectrum [142–144] or sequence-specific [145] [146] antimicrobial activity. Additionally, genome engineering and tools such as CRISPR-Cas [147] and assembly methods including Gibson [148] and yeast assembly [138] should prove useful for the development of new phages with augmented capabilities in modulating microbial communities.

1.3   Thesis Organization and Scope

In the following chapters, I will describe the contributions I have made to the field of microbiome engineering during my tenure in Prof. Timothy Lu's laboratory at MIT. The overarching theme of my thesis is the development of genetic technologies to enable manipulation of the gut microbiota. It is my hope that these technologies will not only prove useful in basic research in the field of microbiome research, but also contribute to devices and cells that may one day reach clinical applications.

In Chapter 2, I describe work in engineering subtractive therapies to selective remove strains from a complex microbial population based on the presence of specific sequences. In the work, published with Robert Citorik as a co-first author in *Nature Biotechnology*, we program CRISPR-Cas nucleases to target antibiotic resistance and virulence genes in *E. coli*. The resultant double-stranded breaks caused by CRISPR-Cas elicit cell death or the loss of the target genes. We explore several different horizontal gene transfer strategies to mediate delivery of genes encoding CRIPSR-Cas into target cells, including transformation, transduction and conjuga-

tion. Bacteriophage-mediated transduction proves to be the most effective means of gene delivery and enables targeted removal of strains from microbial consortia. The work demonstrates the utility of sequence-specific nucleases as programmable narrow-spectrum antimicrobials. Moreover, it highlights the potential of bacteriophage as engineerable targeted antimicrobials in the context of the microbiota.

In Chapter 3, I describe work in developing genetic methodologies that enable synthetic biology in *Bacteroides thetaiotaomicron*. In the work, published with Alex Tucker as a co-first author in *Cell Systems*, we create a suite of genetic parts for precise gene expression in the commensal organism. The majority of the current efforts in microbial engineering for therapeutic purposes employ lactic acid bacteria or *E. coli* as chassis. However, in contrast to these microbes, *Bacteroides* spp. are ubiquitous in human microbiomes, orders of magnitude more abundant, and show greater strain-level stability in microbiomes over time. In order to render *Bacteroides* as genetically malleable as *E. coli*, we developed libraries of constitutive and inducible promoters and ribosome-binding-sites to achieve variable levels of gene expression in *Bacteroides*. Additionally, we demonstrate molecular recording through recombinase-based DNA modification in *B. thetaiotaomicron* as well as re-programmable repression of synthetic and endogenous genes with CRIPSRi. Finally, we validate functionality of these systems in the context of the murine microbiota, suggesting these systems could be used in modulating gene expression *in situ*.

In Chapter 4, I progress the *Bacteroides* engineering work described in Chapter 3 towards a therapeutic goal. First, we characterize the molecular determinants of a major protein secretion pathway in *Bacteroides* spp. *Bacteroides* package numerous proteins with extracellular function into outer membrane vesicles (OMVs) that are excreted by the cell. Membrane-bound lipoproteins are enriched in OMVs and we describe a set of secretion tags that direct *Bacteroides* lipoproteins into the cell membrane. We employ these *Bacteroides* Vesicle Incorporation Tags (BVITs) to package model antigens into *Bacteroides* OMVs and demonstrate their utility in studying immune responses to commensal antigens. Together, the work described in Chapter 3 and Chapter 4 outlines the domestication of an important commensal: developing of

genetic tools for manipulation, characterization of protein secretion, and synthesis of these tools and knowledge to power basic research in the microbiota.

In Chapter 5, I describe a medical device, comprised of electronic and cellular components, that pushes biosensors closer to clinical applications. In this work, performed with Dr. Phillip Nadeau, we create a hybrid-bio-electronic device that can measure the presence or absence of biomarkers in the gastrointestinal tract. As a proof-of-concept, I designed a probiotic *E. coli* biosensor to detect heme and validate the sensor in a mouse model of upper gastrointestinal bleeding. Using custom microelectronics, Dr. Nadeau designed a miniaturized luminometer that can communicate data wirelessly to cellular phones or laptops. When combined, the hybrid device can readily detect the presence of blood *in vitro* and in a porcine model of gastric bleeding. Since the device is modular, biosensor bacteria developed by other groups were easily incorporated into the device for the detection of additional analytes. This work serves as a case study of progressing synthetic biology diagnostic approaches towards real-world application.

Finally, in the concluding chapter, I will describe the major contributions of the preceding chapters to the field of microbiome engineering. I will also outline major challenges faced by the field as well as future directions of the work described herein.

Chapter 2

Sequence-Specific Antimicrobials

The following chapter is adapted from the 2014 Nature Biotechnology publication to which Robert J. Citorik and I contributed equally as first authors.[146]

Abstract

Current antimicrobial strategies tend to be broad-spectrum in nature, leading to indiscriminate killing of commensal bacteria and accelerated evolution of drug resistance. In contrast, we propose a class of programmable-spectrum antimicrobials whose activity can be rationally customized against specific DNA sequences. Using CRISPR-Cas technology, we create RNA-guided nucleases (RGNs) and show that highly efficient, autonomous delivery to microbial populations can be achieved through the use of bacteriophage or conjugation-based vehicles. We design RGNs that selectively target undesirable genes or polymorphisms, including antibiotic resistance and virulence determinants in carbapenem-resistant Enterobacteriaceae and enterohemorrhagic *Escherichia coli*. Delivery of RGNs *in vivo* significantly improves survival in a *Galleria mellonella* infection model. Additionally, RGNs enable sculpting of complex bacterial populations by selective knockdown of targeted strains based on genetic signatures. RGNs constitute a class of highly discriminatory, customizable antimicrobials that enact selective pressure at the DNA level to reduce the prevalence of undesired genes, minimize off-target effects, and enable programmable remodeling

of microbiota.

2.1 Introduction

The emergence and proliferation of multidrug-resistant bacterial pathogens has elicited mounting concern over the dwindling treatment options for these organisms. Recently, carbapenem-resistant Enterobacteriaceae (CRE), a group of intestinal-dwelling Gram-negative bacteria known to cause life-threatening opportunistic infections, were highlighted as one of three most urgent threats among antibiotic-resistant bacteria [149]. Carbapenems have traditionally been reserved as a last resort treatment for Gram-negative infections, but the spread of extended-spectrum β-lactamases has necessitated the increased usage of carbapenems and favored the emergence of carbapenem-resistant strains refractory towards most or all current treatment options. The responsible enzymes, including New-Delhi metallo-β-lactamase 1 (NDM-1), may confer pan-resistance to β-lactam antibiotics and are frequently co-harbored with additional resistance determinants on mobile plasmids that facilitate rapid dissemination within and beyond Enterobacteriaceae [150]. The diversity of multidrug-resistant bacteria compounds the difficulty of developing conventional treatments that target pathogens and commensal reservoirs, but avoid non-specific broad-spectrum activity and undesired pressure on non-target cells.

Here, we introduce an alternative antimicrobial approach that imposes direct evolutionary pressure at the gene level by utilizing efficiently delivered programmable RNA-guided nucleases (RGNs). We engineered the clustered, regularly interspaced, short palindromic repeats (CRISPR)-CRISPR associated (Cas) system, naturally employed in bacteria as a defense strategy against mobile elements [151, 152], to effect cell death or plasmid loss upon detection of genetic signatures associated with virulence or antibiotic resistance. The Type II CRISPR-Cas system of *Streptococcus pyogenes* is an effective, programmable tool for genome editing and gene expression in a wide variety of organisms [153]. The specificity of CRISPR-Cas is dictated by short, spacer sequences flanked by direct repeats encoded in the CRISPR locus, which are

transcribed and processed into mature guide RNAs (crRNA) [154]. With the aid of a trans-activating small RNA (tracrRNA), crRNAs license the Cas9 endonuclease to introduce double-stranded breaks in target DNA sequences [154, 155]. Through simple modifications of spacers in the CRISPR locus, an RGN can direct cleavage of almost any DNA sequence, with the only design restriction being a requisite NGG motif immediately 3' of the target sequence [155]. By packaging RGNs into bacteriophage particles or harnessing mobilizable plasmids, we implemented conditional-lethality devices with high specificity, modularity and multiplexability against undesired DNA sequences (Fig. 2-1).
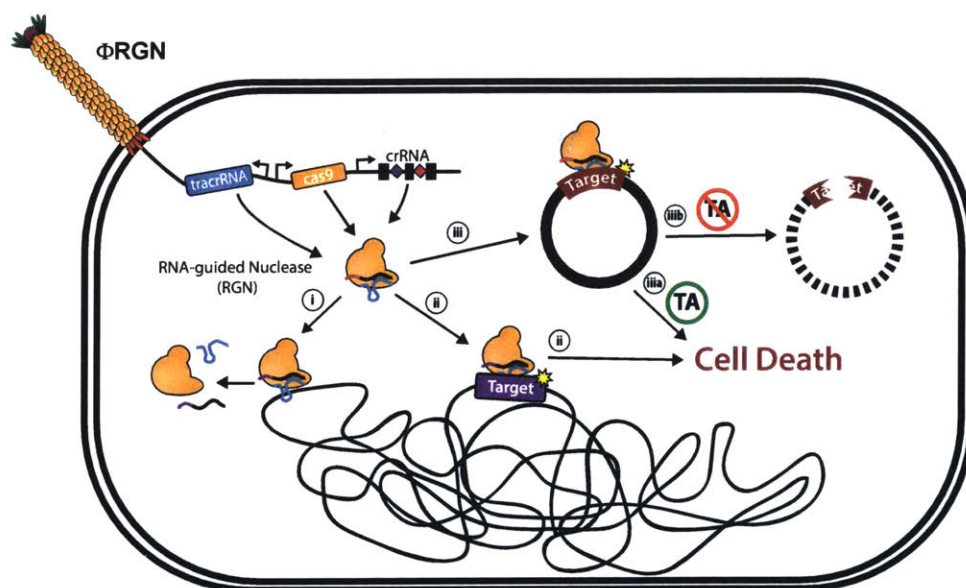
Figure 2-1: **Overview Schematic of CRIPSR-Phage Strategy**: Bacteriophage-delivered RGN constructs differentially affect host cell physiology in a sequence-dependent manner. If the target sequence is: (i) absent, the RGN exerts no effect; (ii) chromosomal, RGN activity is cytotoxic (Fig. 2-5); (iii) episomal, the RGN leads to either (iiia) cell death (Fig. 2-5) or (iiib) plasmid loss (Fig. 2-8A), depending on the presence or absence of toxin-antitoxin (TA) systems (Fig. 2-8B), respectively.

2.2 Transformation

To establish RGN functionality in mediating sequence-specific cytotoxicity, we designed RGNs to induce double-stranded breaks in bla_{SHV-18} or bla_{NDM-1}, which

encode extended-spectrum and pan-resistance to β-lactam antibiotics, respectively [156, 157]. Transformation of plasmid-borne RGNs (pRGNs) into *E. coli* containing a chromosomal copy of these target genes resulted in nearly a thousand-fold decrease in transformation efficiency as compared to wild-type cells lacking the target (Fig. 2-2A). These results corroborate the mutual exclusivity between a functional crRNA and a cognate target locus [158, 159]. Sequence analysis of 30 escape mutants, cells that receive and maintain an RGN plasmid despite the presence of a target sequence, revealed that tolerance was exclusively due to a defective construct, frequently resulting from a spacer deletion within the CRISPR locus (Fig. 2-3). Furthermore, deletion of the tracrRNA as well as inactivation of the RuvC-like nuclease domain of $Cas9_{(D10A)}$ [155] abrogated the loss of transformation efficiency in cells that harboured a target sequence. Thus, a catalytically active endonuclease, tracrRNA and crRNA are necessary and sufficient to mediate sequence-specific cytotoxicity in *E. coli* (Fig. 2-1).

Antibiotic resistance genes often reside on large, multi-copy plasmids capable of autonomous transfer in microbial populations, leading to horizontal dissemination of drug resistance [150]. RGN activity against high-copy plasmids was verified with a GFP-expressing, ColE1-derived vector containing a standard β-lactamase selectable marker (pZE-bla_z-*gfp*) [160] or bla_{NDM-1} (pZE-bla_{NDM-1}-*gfp*). Vectors bearing this ColE1 origin are reported to be present at copy numbers of 50–70 per cell [160]. Transformation of $pRGN_{NDM-1}$, a plasmid-borne RGN targeting bla_{NDM-1}, into cells containing pZE-bla_{NDM-1}-*gfp* led to a three-$\log_{10}$ reduction in transformants retaining carbenicillin resistance, whereas transformation of $pRGN_{NDM-1}$ into cells containing target-free pZE-bla_z-*gfp* did not lead to a reduction in resistant transformants (Fig. 2-2B). The activity of RGNs is therefore sufficient to exclude even high-copy antibiotic resistance plasmids from cells and can re-sensitize a resistant population to antibiotics. Similarly, transformation of $pRGN_{NDM-1}$ into cells possessing pZE-bla_{NDM-1}-*gfp* led to an approximately thousand-fold decrease in GFP-expressing cells, as measured by flow cytometry, but no decrease was found with transformation of $pRGN_{NDM-1}$ into cells possessing pZE-bla_z-*gfp* (Fig. 2-2C).
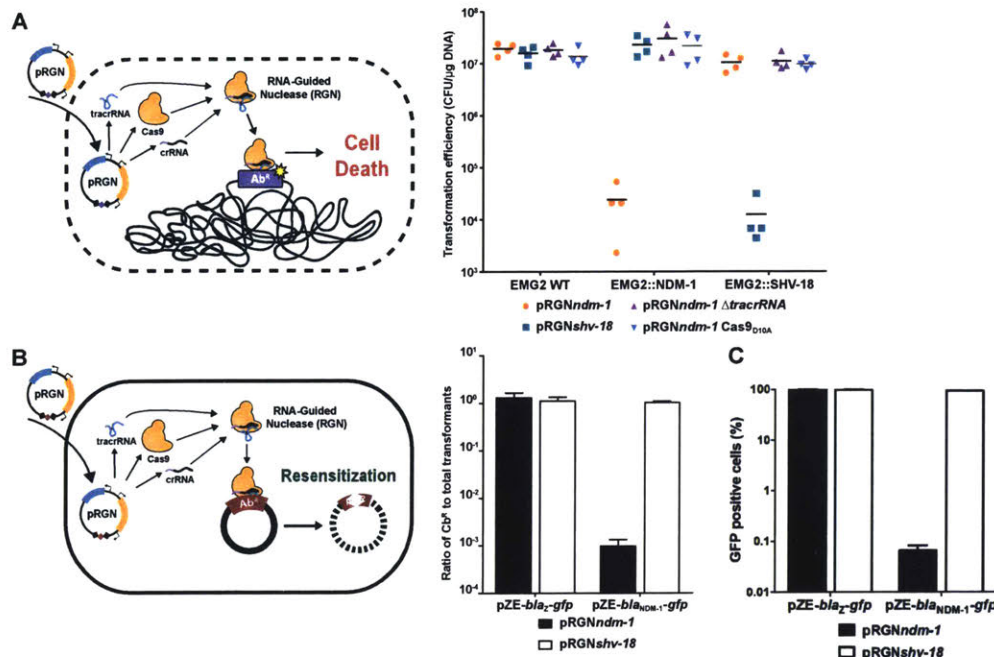
Figure 2-2: **Design and validation of programmable RGN constructs by transformation**: (A) Plasmids $pRGN_{NDM-1}$, $pRGN_{SHV-18}$, $pRGN_{NDM-1}$ ΔtracrRNA, and $pRGN_{NDM-1}$ Cas_{D10A} were transformed into competent wild-type EMG2 (EMG2 WT) as well as otherwise isogenic strains containing chromosomally integrated bla_{NDM-1} (EMG2::NDM-1) or bla_{SHV-18} (EMG2::SHV-18). Transformants were enumerated on LB+chloramphenicol (Cm) to select for pRGN transformants and to determine transformation efficiencies, which demonstrated the specific incompatibility of an RGN construct and its cognate protospacer (n=4). (B) Plasmids $pRGN_{NDM-1}$ and $pRGN_{SHV-18}$ were transformed into EMG2 cells containing either pZE-bla_{NDM-1}-gfp or pZE-bla_Z-gfp plasmids. Transformants, first selected in appropriate antibiotic media, were enumerated on LB+Cm or LB+Cm+carbenicillin (Cb) agar to calculate the ratio of transformants retaining Cb resistance (Cb^R) to total transformants. Error bars indicate s.e.m. of three independent experiments, each with three biological replicates (n=9). (C) EMG2 cells containing either pZE-bla_Z-gfp or pZE-bla_{NDM-1}-gfp plasmids were transformed with $pRGN_{NDM-1}$ or $pRGN_{SHV-18}$ plasmids and transformants were selected overnight in LB+chloramphenicol (no antibiotic selection for plasmid maintenance was applied). Plasmid loss was determined by calculating the percentage of GFP-positive cells following gating by forward and side scatter. Error bars indicate s.e.m. of measurements from three independent experiments, each with three biological replicates (n=9).

Mutation	Number	%
Deletion of spacer and one repeat	19	63.3
tracrRNA Insertion (A at nt140/141)	6	20.0
Transposon Insertion (Tn10.10)	5	16.7
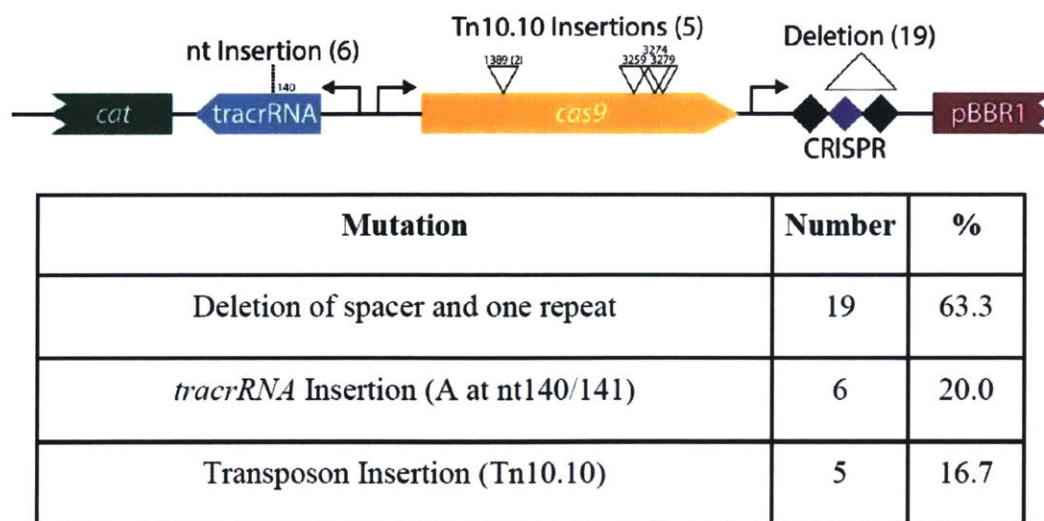
Figure 2-3: **Characterization of escape mutants that tolerated transformation of a cytotoxic RGN construct**: EMG2::NDM-1 or EMG2::SHV-18 colonies that tolerated transformation of the pRGN$_{NDM-1}$ or pRGN$_{SHV-18}$ plasmids (Fig. 2-2B) were re-isolated and sequenced to identify escape mutations. Spacer deletion in the CRISPR locus, point mutations in *tracrRNA* and transposon insertions in *cas9* led to pRGN inactivation in successful transformants. Five escape mutants from three independent experiments were sequenced per strain (n=30).

2.3 Conjugation

The viability of RGNs for antimicrobial therapy hinges on high-efficiency delivery of genetic constructs to bacterial cells. We explored two mechanisms of horizontal gene transfer naturally employed by bacteria to acquire foreign genetic elements: plasmid conjugation and viral transduction. Although constrained by requirements for cell-cell contact, conjugative plasmids often possess wide host ranges and no recipient factors necessary for DNA uptake have been identified [161]. Efficient transfer of RGNs was achieved using the broad-host-range plasmid R1162 mobilized by *E. coli* S17-1, which contains the conjugative machinery of plasmid RP4 integrated into its chromosome. In filter mating experiments, conjugative transfer of RGNs elicited a 40–60-fold reduction in target carbenicillin-resistant recipient cells (Fig. 2-4B). Under selection for transconjugants, transfer of RGNs into recipients yielded a 2–3-$\log_{10}$ reduction in target cells as compared to controls, suggesting that conjugation effi-

ciency, as opposed to RGN activity, limits RGN efficacy in this context (Fig. 2-4C). Future work may be necessary to further optimize the efficiency of conjugation-based delivery vehicles for antimicrobials based on mobilizable RGNs.
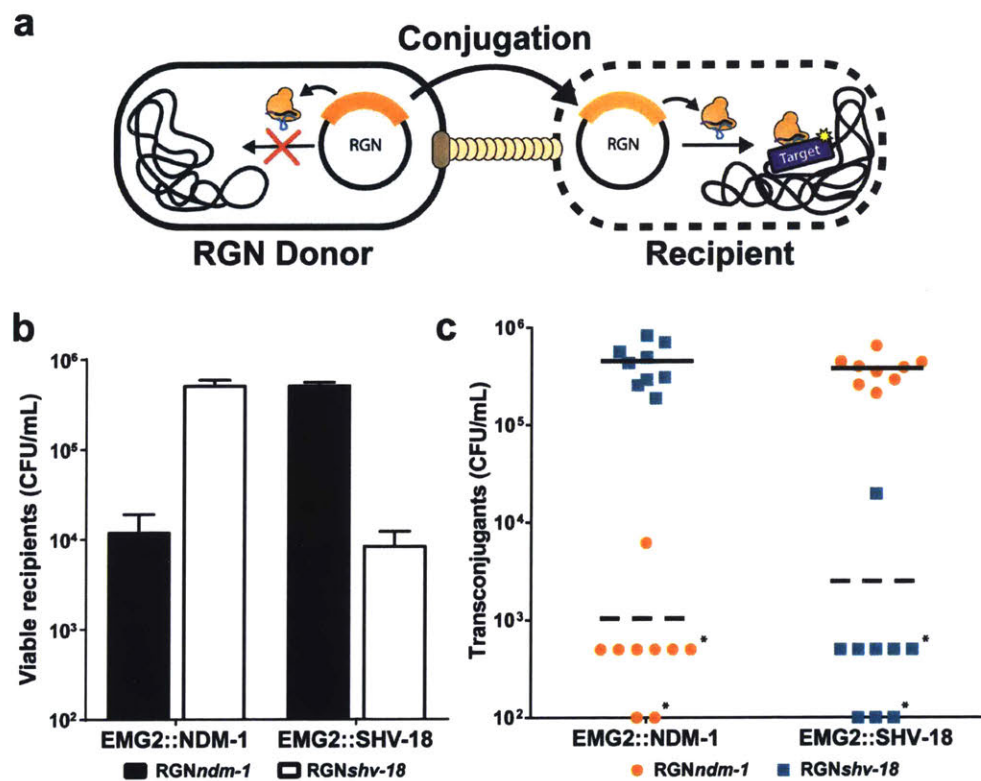
Figure 2-4: **Mobilizable RGNs can be conjugated into target bacteria for selective removal of multidrug resistance**: (A) Schematic of mobilizable RGN-mediate cell killing. (B-C) S17-1 λpir donor cells possessing RGN_{NDM-1} or RGN_{SHV-18} were mated at a donor:recipient ratio of $340\pm66:1$ for 3 hours with EMG2 recipient cells that contain bla_{NDM-1} (EMG2::NDM-1) or bla_{SHV-18} (EMG2::SHV-18) integrated into the chromosome. Cultures were plated on LB+carbenicillin to select for surviving recipient cells (B) and LB+chloramphenicol+carbenicillin to select for transconjugants (C) (chloramphenicol resistance is encoded by the RGN plasmids). (C) Transfer of a mobilizable RGN into cells containing the cognate target sequence (dashed line) reduced the number of viable transconjugants to the limit of detection (∗) (100 CFU/mL or 500 CFU/mL for three or six of the biological replicates, respectively) in almost all cases. Error bars indicate s.e.m. of three independent experiments, each with three biological replicates (n=9).

2.4 Transduction

Bacteriophages are natural predators of prokaryotes and are highly adept at injecting DNA into host cells. To implement phage for RGN delivery, we engineered phagemid vectors by pairing RGN constructs targeting bla_{NDM-1} or bla_{SHV-18} with an f1 origin for packaging into M13 particles. Phage-packaged RGN_{NDM-1} (ΦRGN_{NDM-1}) was capable of comprehensively transducing a population of E. coli EMG2 (Fig. 2-5.A). To test the $\Phi RGNs$, we conjugated native plasmids containing bla_{NDM-1} (pNDM-1) or bla_{SHV-18} (pSHV-18) from clinical isolates into EMG2. Treatment of the EMG2 pNDM-1 or EMG2 pSHV-18 strains with the cognate $\Phi RGNs$ resulted in 2–3-$\log_{10}$ reductions in viable cells even in the absence of any selection (Fig. 2-5.B). Furthermore, $\Phi RGNs$ engendered no toxicity against wild-type EMG2 or EMG2 containing non-cognate plasmids (Fig. 2-5.B).

In naturally occurring Type II CRISPR-Cas systems, the CRISPR locus may contain multiple spacers, each of which is processed into independent crRNA molecules that license Cas9 to cleave cognate DNA sequences [154]. To explore the utility of a single ΦRGN exhibiting activity against more than one genetic signature, we engineered a construct containing two spacers encoding two different crRNAs for targeting the bla_{NDM-1} and bla_{SHV-18} resistance genes ($\Phi RGN_{NDM-1/SHV-18}$). $\Phi RGN_{NDM-1/SHV-18}$ generated 2–3-$\log_{10}$ reductions in viable cells counts of EMG2 pNDM-1 or EMG2 pSHV-18, but not of wild-type EMG2 (Fig. 2-5.B). Thus, RGNs may be multiplexed against different genetic signatures, enabling simultaneous targeting of a variety of virulence factors and resistance genes that may exist in microbial populations.

In addition to antibiotic-modifying enzymes, such as β-lactamases, alterations in host proteins constitute a major antibiotic resistance mechanism [162]. Owing to the specificity of the CRISPR-Cas system in prokaryotes, we suspected RGNs could discriminate between susceptible and resistant strains that differ by a single nucleotide mutation in DNA gyrase ($gyrA$), which confers resistance to quinolone antibiotics [162]. Indeed, $\Phi RGN_{gyrA_{D87G}}$ exhibited specific cytotoxicity only towards quinolone-resistant E. coli harbouring the chromosomal $gyrA_{D87G}$ mutation and not
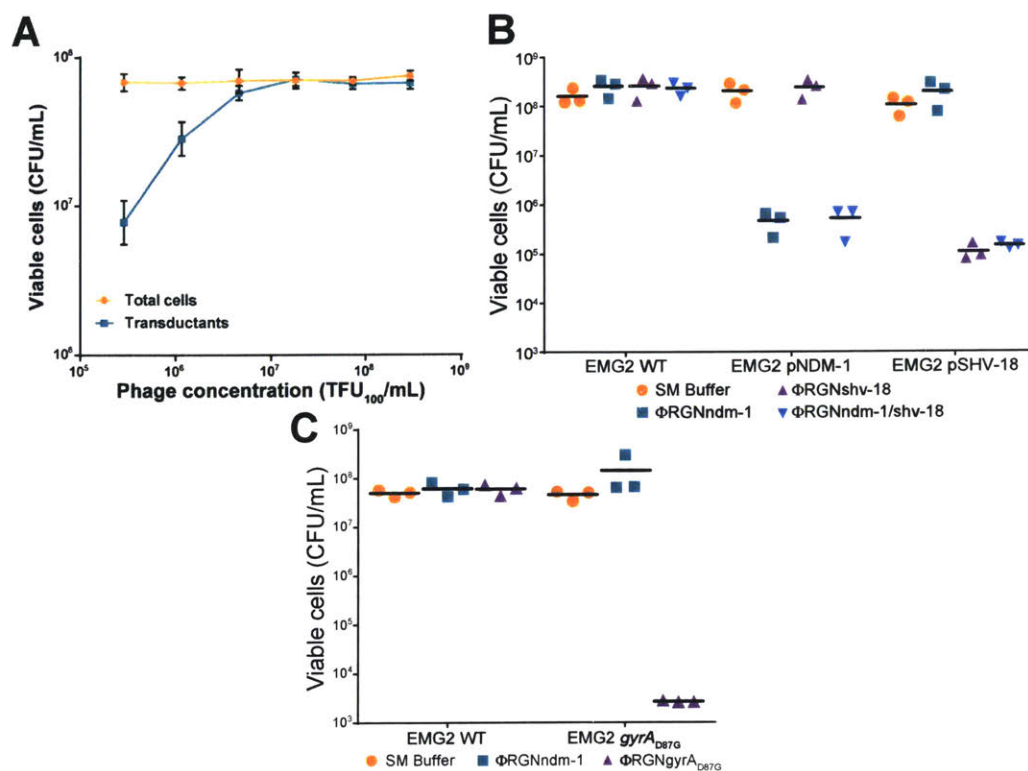
Figure 2-5: **RGN constructs delivered via bacteriophage particles (ΦRGN) exhibit efficient and specific antimicrobial effects against strains harbouring plasmid or chromosomal target sequences**: (A) Approximately 6.5×10^6 CFU/mL of EMG2 wild-type cells were incubated with dilutions of ΦRGN_{NDM-1} phagemid for 2 hours and plated onto LB and LB+kanamycin to determine the highest dilution of the purified ΦRGN stock able to transduce approximately 100% of the recipient cell population (defined as TFU_{100}/mL). The ΦRGN_{NDM-1} phagemid encodes a kanamycin resistance gene. Error bars indicate s.e.m. of three independent experiments (n=3). (B) Treatment of EMG2 wild-type (WT) or EMG2 containing native resistance plasmids, pNDM-1 (encoding bla_{NDM-1}) or pSHV-18 (encoding bla_{SHV-18}), with SM buffer, ΦRGN_{NDM-1}, ΦRGN_{SHV-18}, or multiplexed $\Phi RGN_{NDM-1/SHV-18}$ at a multiplicity of infection (MOI) ~20 showed sequence-dependent cytotoxicity as evidenced by a strain-specific reduction in viable cell counts (n=3). (C) *E. coli* EMG2 WT or EMG2 $gyrA_{D87G}$ populations were treated with SM buffer, ΦRGN_{NDM-1} or $\Phi RGN_{gyrA_{D87G}}$ at MOI ~20, and viable cells were determined by plating onto LB agar (n=3).

towards otherwise isogenic strains with the wild-type *gyrA* gene (Fig. 2-5.C).

Killing curves revealed that ΦRGNs mediated rapid killing of target cells, with viable cell counts that decayed exponentially ($t_{1/2}$ ~13 minutes) and maximal bactericidal effect achieved by 2–4 hours (Fig. 2-6.A). Moreover, ΦRGN antimicrobial

activity increased with phagemid particle concentration (Fig. 2-6.B).
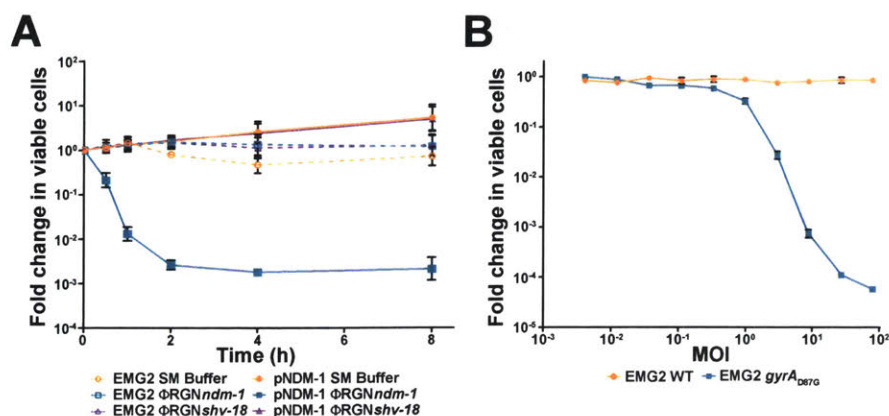
Figure 2-6: **Characterization of ΦRGN-mediated killing of antibiotic-resistant bacteria**: (A) Time-course treatment of EMG2 WT or EMG2 pNDM-1 with SM buffer, ΦRGN_{NDM-1} or ΦRGN_{SHV-18} at a multiplicity of infection (MOI) ∼20. Data represent the fold change in viable colonies at indicated time points relative to time 0h. (B) Dose response curve of EMG2 WT and EMG2 $gyrA_{D87G}$ treated with various concentrations of $\Phi RGN_{gyrA_{D87G}}$ for two hours. Data represent fold change in viable colonies relative to SM buffer treated samples. Error bars represent s.e.m. of three independent biological replicates (n=3).

To further characterize the cellular response to RGN-mediated targeting, we assessed treatment of cells possessing a GFP reporter under SOS regulation. *E. coli* and other bacteria respond to chromosomal double-stranded breaks, including those artificially generated by the meganuclease I-SceI, by inducing DNA repair through the activation of the SOS response [163]. We observed a 2.6- or 4.0-fold increase in fluorescence in cells containing the reporter plasmid and a plasmid-borne (bla_{NDM-1}) or chromosomal ($gyrA_{D87G}$) target site, respectively, when treated with the cognate versus non-cognate ΦRGNs (Fig. 2-7). These results confirm that RGNs can induce DNA damage in target cells and demonstrate that they can be coupled with SOS-based reporters to detect specific genes or sequences, even at the single-nucleotide level.
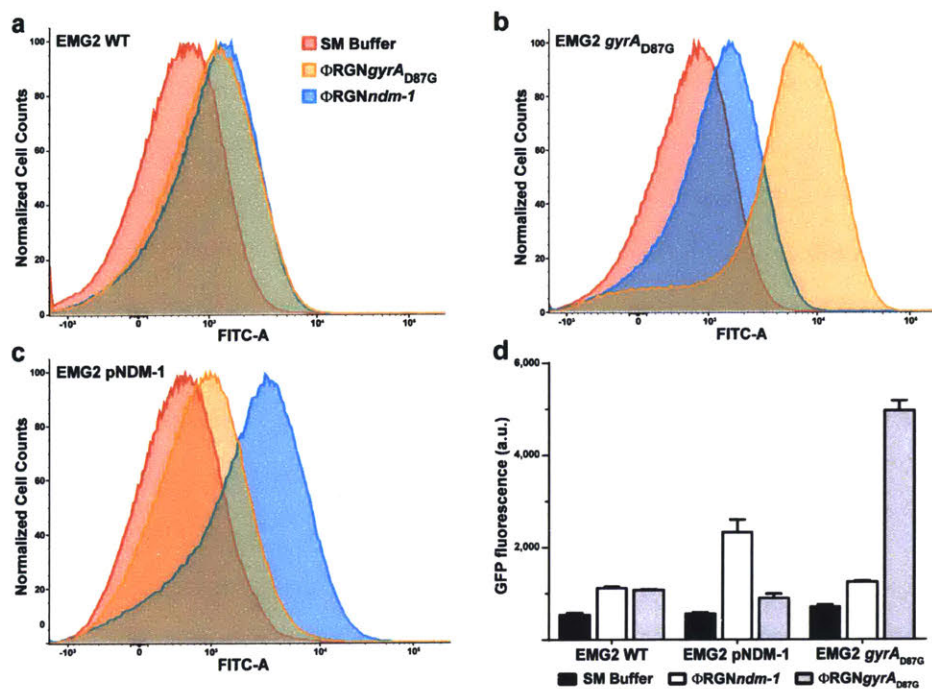
Figure 2-7: **Treatment of *E. coli* with ΦRGNs induces DNA damage and an SOS response in cells that possess a cognate target sequence**: EMG2 wild-type (WT) (A), EMG2 *gyrA*$_{D87G}$ (B) and EMG2 pNDM-1 (C) containing the pZA3LG reporter plasmid were treated with either SM buffer, ΦRGN$_{NDM-1}$ or ΦRGN$_{gyrA_{D87G}}$. GFP expression on pZA3LG is under the control of the SOS-responsive P$_{L(LexO)}$ promoter [164]. Injection of single-stranded phagemid DNA led to a mild induction of the SOS response, whereas RGN activity in cells that possessed a cognate target sequence led to stronger induction of SOS. Histograms were generated by combining data from four biological replicates and are normalized to the mode of the population. (D) Summary of flow cytometry histograms. The arithmetic means of the geometric mean fluorescence of populations in A–C were calculated across four independent biological replicates (n=4). Error bars represent s.e.m.

2.4.1 Toxin-Antitoxin Activation

We were intrigued to observe that targeted cleavage of *bla*$_{NDM-1}$ with ΦRGN$_{NDM-1}$ in the context of the native plasmid was lethal to host cells, whereas targeted cleavage of the same gene in a standard cloning vector was not ('pNDM-1' versus 'pZA-*ndm1-gfp*', respectively, in Fig. 2-8A). Therefore, we hypothesized that ΦRGN-induced plasmid loss in itself does not elicit lethality, but rather results in cytotoxicity via other co-harboured plasmid-borne functions. Toxin-antitoxin systems are components of natural plasmids that ensure persistence in bacterial populations by inhibit-

ing growth of daughter cells that fail to inherit episomes following cell division. These addiction modules traditionally consist of a labile antitoxin that quenches the activity of a stable toxin. Owing to the differential stability of these two components, cessation of gene expression upon plasmid loss leads to depletion of the antitoxin pool faster than the toxin pool, resulting in de-repression of toxin activity and, ultimately, stasis or programmed cell death [165]. Analysis of the sequenced pSHV-18 plasmid revealed the presence of a unique toxin-antitoxin module, *pemIK*, which is commonly found among isolates harbouring extended-spectrum β-lactamases [166]. When complemented with the PemI antitoxin expressed constitutively in trans (pZA31-*pemI*), ΦRGN_{SHV-18} treatment of EMG2 pSHV-18 abrogated cytotoxicity and instead resulted in resensitization of this multidrug-resistant strain to carbenicillin (Fig. 2-8B). We attributed this effect to PemI inactivation of the PemK toxin, thus enabling loss of the pSHV-18 plasmid without concomitant bacterial killing. Addiction modules can therefore dictate the outcome of ΦRGN activity on episomal targets, as their presence leads to cytotoxicity and their absence or neutralization to plasmid loss.
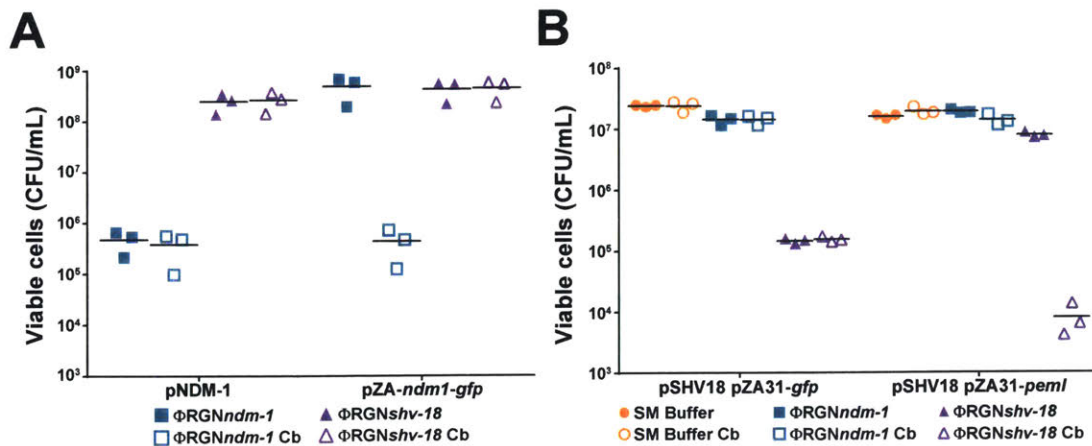
Figure 2-8: **RGN-mediated targeting of toxin-antitoxin systems can lead to cytotoxicity**: (A)EMG2 *E. coli* containing the natural pNDM-1 plasmid or the *bla*NDM-1 gene in a synthetic expression vector (pZA-*ndm1-gfp*) were treated with either ΦRGN_{NDM-1} or ΦRGN_{SHV-18} at MOI ~20 and plated onto both non-selective LB and LB+carbenicillin (Cb) to select for *bla*NDM-1-containing cells. ΦRGN_{NDM-1} treatment of cells harboring pNDM-1 resulted in a reduction in viability in the absence of selection, whereas ΦRGN_{NDM-1} treatment of cells with pZA-*ndm1-gfp* demonstrated similar cytotoxicity only under selective pressure for maintenance of the pZA-*ndm1-gfp* plasmid. (B) EMG2 pSHV-18 complemented with the cognate antitoxin (pZA31-*pemI*) for the PemK toxin or a control vector (pZA31-*gfp*) was treated with SM buffer, ΦRGN_{NDM-1} or ΦRGN_{SHV-18}. Cultures were plated on LB and LB+Cb and colonies were enumerated to assess cytotoxicity or plasmid loss.

2.5 Targeting Virulence Genes in *E. coli* O157:H7

To further demonstrate the versatility of RGNs for specifically combating pathogens, we designed a ΦRGN to target intimin, a chromosomally encoded virulence factor of enterohemorrhagic *E. coli* O157:H7 (EHEC) necessary for intestinal colonization and pathology. Encoded by the *eae* gene, intimin is a cell-surface adhesin that mediates intimate attachment to the host epithelium, permitting subsequent disruption of intestinal tight junctions and effacement of microvilli [167]. Treatment of EHEC with ΦRGN_{eae} resulted in a 20-fold reduction in viable cell counts; this cytotoxicity was increased an additional 100-fold under kanamycin selection for ΦRGN_{eae} transductants (Fig. 2-9A). The increase in cytotoxicity with selection for cells receiving the construct implies that the efficacy of ΦRGN treatment was limited by delivery in this strain. Furthermore, ΦRGN treatment was assessed in *Galleria mellonella* lar-

vae, an infection model that yields virulence data often predictive for higher-order mammals [168]. This model has also been used to evaluate the efficacy of antimicrobials or phage therapy against various Gram-negative, Gram-positive and fungal pathogens [168]. Administration of ΦRGN to EHEC-infected *G. mellonella* larvae significantly improved survival over no treatment or an off-target ΦRGN control (Log-rank test, $p<0.001$) (Fig. 2-9B). Moreover, ΦRGN_{eae} was significantly more effective than chloramphenicol treatment, to which the EHEC strain was resistant (Log-rank test, $p<0.05$). Although ΦRGN_{eae} treatment was inferior to carbenicillin, to which the bacteria were susceptible, these data support RGNs as viable alternatives for cases where bacteria are highly resistant to existing antibiotics. Improvements in delivery efficiency with ΦRGN_{EAE} would be expected to improve treatment efficacy and outcome.
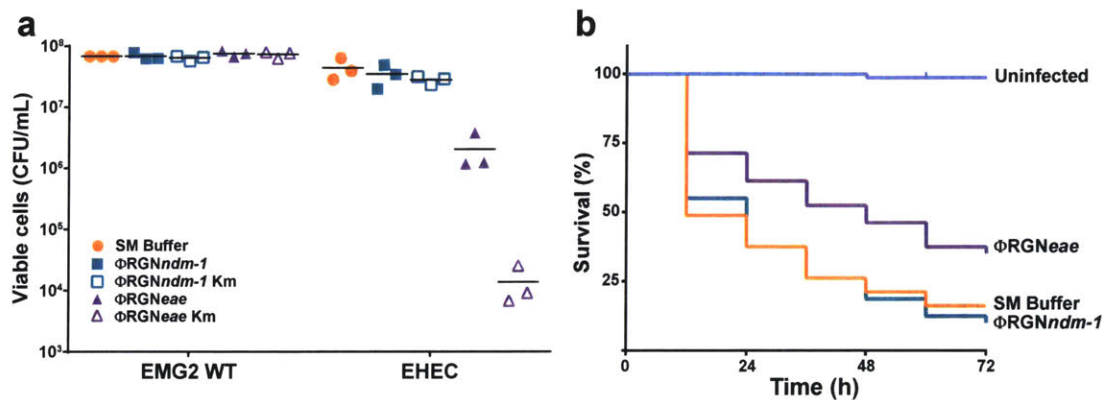
Figure 2-9: **ΦRGN particles elicit sequence-specific toxicity against enterohemorrhagic *E. coli in vitro* and *in vivo*:** (A)*E. coli* EMG2 wild-type (WT) cells or ATCC 43888 F' (EHEC) cells were treated with SM buffer, ΦRGN_{NDM-1} or ΦRGN_{eae} at a multiplicity of infection (MOI) ~100 and plated onto LB agar to enumerate total cell number or LB+kanamycin (Km) to select for transductants with ΦRGNs (n=3). (b) *Galleria mellonella* larvae were injected with either PBS or approximately $4x10^5$ colony forming units (CFU) of EHEC. Subsequent administration of ΦRGN_{eae} at MOI ~30 significantly improved survival compared to SM buffer or ΦRGN_{NDM-1} treatment (Log-rank test, p¡0.001). Survival curves represent an aggregate of four independent experiments, each with 20 worms per treatment group (n=80).

2.6 Population Sculpting

In addition to implementing targeted antimicrobial therapies, RGNs can be used to sculpt the composition of complex bacterial populations (Fig. 2-10). Current therapies that use a prebiotic, probiotic or drug to modify the human microbiota have demonstrated potential for alleviating various disease states, but remain poorly characterized in terms of off-target effects and the specific mechanisms by which they act [169]. In concert with the host range of the delivery vehicle, RGN activity can selectively remove bacteria with specific genomic content. This could reduce the prevalence of unwanted genes, including antibiotic resistance and virulence loci, or metabolic pathways from bacterial communities without affecting bystanders.

To demonstrate a proof-of-principle for 'bacterial knockdowns' using RGNs, we constructed a synthetic consortium comprised of three phage-susceptible *E. coli* strains with differential antibiotic resistance profiles. We used β-lactam-resistant *E. coli* EMG2 pNDM-1, quinolone-resistant RFS289 ($gyrA_{D87G}$), and chloramphenicol-resistant CJ236. Application of ΦRGN$_{NDM-1}$ elicited >400-fold killing of EMG2 pNDM-1, while leaving RFS289 and CJ236 cell populations intact. Treatment with ΦRGN$_{gyrA_{D87G}}$ resulted in >20,000-fold killing of RFS289 without a concomitant reduction in EMG2 pNDM-1 or CJ236 (Fig. 2-10). These results demonstrate that RGNs can selectively knockdown bacteria that contain target DNA sequences while allowing the remaining non-target bacteria to dominate the population. Adapting this approach for tuning endogenous microbiota could be accomplished by delivering RGNs *in vivo* via broad-host-range phages or phage cocktails, or with conjugative plasmids. An appropriately targeted bacterial knockdown approach could be employed in functional studies of complex microbiota and to complement additive therapies, such as probiotics, for microbiome-associated diseases by clearing specific niches or removing defined genes from bacterial populations.
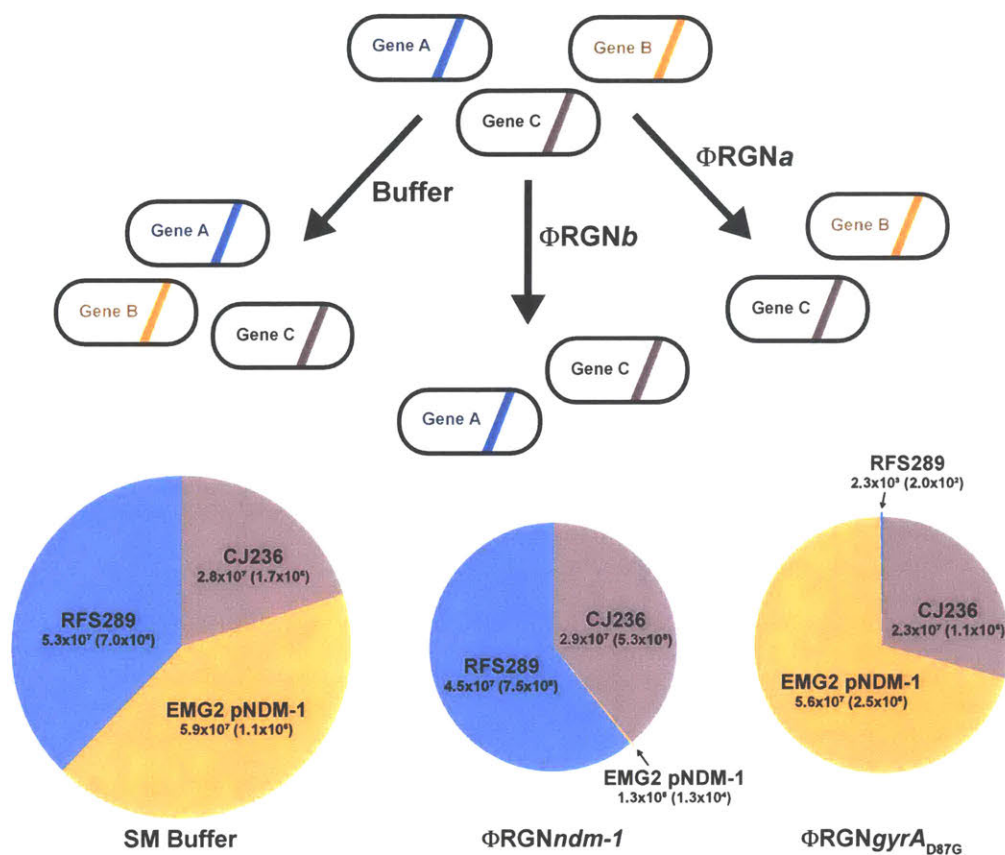
Figure 2-10: **Programmable remodeling of a synthetic microbial consortium**: A synthetic population composed of three different *E. coli* strains was treated with either SM buffer, ΦRGN_{NDM-1}, or $\Phi RGN_{gyrA_{D87G}}$ at an MOI ~ 100 and plated onto LB with chloramphenicol, streptomycin or ofloxacin to enumerate viable cells of *E. coli* CJ236, EMG2 pNDM-1 or RFS289 strains, respectively. ΦRGN_{NDM-1} targets bla_{NDM-1} in EMG2 pNDM-1 and $\Phi RGN_{gyrA_{D87G}}$ targets the $gyrA_{D87G}$ allele in RFS289. Circle area is proportional to total population size and numbers represent viable cell concentrations (CFU/mL) of each strain after the indicated treatment. The s.e.m. based on three independent experiments is indicated in parentheses (n=3).

2.7 Discussion

In light of the rising tide of antibiotic resistance, interest in engineered cellular and viral therapeutics as potential biological solutions to infectious disease has resurged. By repurposing parts developed by nature, synthetic biologists have designed artificial gene circuits for antimalarial production [170], engineered probiotics [171] and phage therapeutics to eradicate biofilms [172] or potentiate antibiotic activity [173, 174]. Here, we demonstrate that transmissible CRISPR-Cas systems can act as a

platform for programmable antimicrobials that harness site-specific cleavage to induce cytotoxicity, activate toxin-antitoxin systems, resensitize bacterial populations to antibiotics, and sculpt bacterial consortia. This work complements the recent finding that the *Vibrio cholerae* phage, ICP1, encodes its own CRISPR-Cas system to counteract a host-encoded phage defense locus [175] and that CRISPR-Cas constructs transformed into electrocompetent cell populations are incompatible with cells that contain cognate target sequences [159, 176, 177]. In contrast to these latter studies, we show that CRISPR-Cas technology can be applied *in situ* for the removal of undesired genes from microbial populations and *in vivo* to treat infection in the absence of artificial selection. Moreover, we demonstrate that RGNs can be used to artificially activate plasmid-borne toxin-antitoxin systems, which has recently become an attractive antimicrobial strategy [178]. In addition to validating antimicrobial activity, we further demonstrate potential applications of RGNs in the deletion of plasmids from cells or the detection of DNA elements with up to single nucleotide resolution using a DNA-damage-responsive reporter.

Since CRISPR-Cas systems are widely conserved in prokaryotes, the development and optimization of novel delivery vehicles will aid in the creation of new RGNs capable of targeting additional strains, including multidrug-resistant pathogens as well as key players in natural microbial communities, such as the human microbiome. Phage-based therapies are dependent on their ability to deliver nucleic acids into bacteria, which can be resisted through a variety of mechanisms [179]. These delivery vehicles can be limited to a subset of bacteria defined by the chosen phage, such that the design of programmable antimicrobials may require additional considerations as to the phage platform chosen. However, rational modification to phage host range through tail fiber alterations [180] or the use of bacteriophage cocktails [181] can mitigate the host range limitations of phage-based therapies. Although the use of bacteriophages in humans has been met with challenges [181], especially in the Western world, a renaissance in phage-based therapeutics has begun to address these challenges, such as demonstrating safety in humans [182], improving the persistence of phages remaining in circulation by reducing their clearance by the host [183] and minimizing

endotoxin release by using non-lytic phage engineered with heterologous kill functions [184]. Additionally, we devised a complementary delivery strategy using a mobilizable broad-host-range system to introduce RGNs to recipient cells via conjugation. The use of conjugative delivery from probiotics into target bacteria would enable a platform where engineered cells could integrate complex environmental cues and execute lethal payload delivery, akin to previously described sentinel cells [185]. Future work is needed to improve the efficiency and spectra of delivery strategies for RGNs, which may include broad-host-range bacteriophages and more efficient conjugative strategies, as well as chemical delivery technologies.

Delivery systems which extend to higher organisms could also enable RGNs to modulate the prevalence of specific genes in wild-type populations [186]. Owing to the modularity and simplicity of CRISPR-Cas engineering, libraries of multiplexed ΦRGNs can be rapidly constructed to simultaneously target a plethora of antibiotic resistance and virulence determinants and to sculpt complex microbial communities. The addition of facile, sequence-informed rational design to a field dominated by time- and cost-intensive screening for broad-spectrum small-molecule antibiotics could have the potential to reinvigorate the dry pipeline of new antimicrobials.

2.8 Experimental Details

Strains and Culture Conditions

Unless otherwise noted, bacterial cultures were grown at 37°C with LB medium (BD Difco). Where indicated, antibiotics were added to the growth medium to the following final concentrations: $100\mu g/mL$ carbenicillin (Cb), $30\mu g/mL$ kanamycin (Km), $25\mu g/mL$ chloramphenicol (Cm), $100\mu g/mL$ streptomycin (Sm), and $150ng/mL$ ofloxacin (Ofx).

Strain Construction

E. coli EMG2 Sm^R was generated by plating an overnight culture of *E. coli* EMG2 onto LB+Sm. Spontaneous resistant mutants were re-streaked onto LB+Sm and an isolated colony was picked and used as the recipient for conjugation of the multidrug resistance plasmids. Overnight cultures of EMG2 Sm^R (recipient), *E. coli* CDC1001728 (donor for pNDM-1) and *K. pneumoniae* K6 (donor for pSHV-18) were washed in sterile PBS and $100\mu L$ of donor and recipient were spotted onto LB agar plates and incubated at 37°C overnight. Transconjugants were harvested by scraping the cells in 1mL of sterile PBS and plating onto LB+Sm+Cb.

The chromosomal integrations of the $bla_{\text{NDM-1}}$ and $bla_{\text{SHV-18}}$ β-lactamase genes and generation of EMG2 $gyrA_{\text{D87G}}$ were performed by λ-Red recombineering using the pSIM9 system [187].Templates for integration at the non-essential *lacZYA* locus were generated by amplifying the $bla_{\text{NDM-1}}$ and $bla_{\text{SHV-18}}$ genes from lysates of CDC1001728 and K6 using the primers rcD77/78 and rcD73/74, respectively. Templates for construction of EMG2 $gyrA_{\text{D87G}}$ were obtained by amplifying *gyrA* from RFS289 using primers mmD155/161.

Plasmid Construction

To generate the RGN plasmids, an intermediate vector pZA-RGN$\varnothing$, which lacks a CRISPR locus, was created. The tracrRNA and $P_{\text{L(TetO-1)}}$ promoter were synthesized

(Genewiz) and amplified using primers mmD98/99, cas9 was amplified from pMJ8067 using primers mmD74/75 and the vector backbone was amplified from pZA11G using primers mmD82/83. Each PCR product was purified and ligated by Gibson assembly [188]. To create the final backbone vector for the RGN plasmids, the pBBR1 origin, chloramphenicol resistance marker, tL17 terminator, and CRISPR locus cloning site were amplified from an intermediate vector pBBR1-MCS1-tL17 using mmD151/154, digested with NheI and SacI-HF, and ligated with pZA-RGN∅ digested with SacI-HF and AvrII to create pZB-RGN∅. Digestion of this vector with PstI-HF and XbaI allowed for the insertion of assembled CRISPR loci. The ΔtracrRNA pRGN$_{NDM-1}$ plasmid was created by amplification of pRGN$_{NDM-1}$ with mmD162/163, ClaI digestion, and self-ligation. The Cas9$_{D10A}$ mutant plasmids were constructed through site-directed mutagenesis of pRGN$_{NDM-1}$ with primers mmD108/109 and the KAPA HiFi PCR kit (KAPA Biosystems).

The CRISPR loci were constructed through isothermal annealing and ligation of short, single-stranded oligonucleotides (Integrated DNA Technologies). Each spacer and repeat piece was built by a corresponding oligo duplex connected to adjacent pieces by 6bp overhangs. In addition, the terminal repeats were designed to contain a 17bp extension comprised of a BsaI restriction site to generate an overhang that allowed insertion into the pUC57-Km-crRNA∅ backbone vector synthesized by Genewiz.

To assemble the CRISPR loci, 500pmol of sense and antisense oligos in a given duplex were annealed by boiling for 10 minutes at 99°C and cooled to room temperature. 300pmol of each annealed duplex were combined with 15U of T4 polynucleotide kinase (Affymetrix), 400U of T4 DNA ligase (NEB), T4 ligase buffer (NEB) and ddH$_2$O to a volume of 20μL. Following incubation at 25°C for 1 hour, the reaction products were purified using a Qiagen QIAquick PCR Purification Kit. Purified products were digested for three hours with BsaI-HF and re-purified using QIAquick. To prepare the crRNA backbone vector, pUC57-Km-crRNA∅ was amplified using primers mmD104/105, subsequently digested with BbsI to generate compatible overhangs, and ligated with the assembled CRISPR loci. Positive clones of the CRISPR

loci were digested from the entry vector using PstI-HF and XbaI and ligated into pZB-RGN∅ digested with the same enzymes to create the final RGN plasmids.

Phagemid vector pZEf-*gfp* was created previously by adding the f1 origin amplified from the yeast shuttle pRS series [189] into pZE22-*gfp*12. The RGN constructs consisting of the genes encoding the tracrRNA, Cas9, and a sequence-targeting crRNA were amplified as a single product from the respective pRGN vectors using KAPA HiFi polymerase (Kapa Biosystems) with primers rcD169/183 and digested with AvrII and XmaI (New England Biolabs). These inserts were ligated with a backbone derived from amplifying the kanamycin resistance cassette, ColE1 replication origin and the f1 origin required for packaging into M13 particles off of pZEf-*gfp* with primers rcD184/185 and digesting with the same enzymes. Ligated plasmids were transformed into *E. coli* DH5αPro for sequence verification and plasmid purification.

The pZE-*bla*$_{NDM-1}$-*gfp* and pZA-*bla*$_{NDM-1}$-*gfp* vectors were constructed by swapping the antibiotic resistance cassette of the Lutz-Bujard vectors pZE12G and pZA12G12. The *bla*$_{NDM-1}$ gene was amplified from a lysate of CDC1001728 using primers mmD8/9 and the PCR product was digested with SacI-HF and XhoI. The digested product was ligated into the Lutz-Bujard vectors digested with the same enzymes.

The PemI antitoxin complementation plasmid pZA31-*pemI* was created by first amplifying the *pemI* coding sequence from pSHV-18 with mmD253/254. The PCR product was digested with BamHI and KpnI and ligated with the large fragment of a pZA31G digest with the same enzymes. The SOS-responsive pZA3LG reporter plasmid was derived from pZE1LG [164] by swapping the origin of replication and antibiotic resistance marker with pZA31G using AatII and AvrII as restriction enzymes.

Mobilizable RGNs were created by first amplifying the R1162 replication origin and *oriT* using mmD266/267. The chloramphenicol selection marker and RGN locus were amplified from pRGN$_{NDM-1}$ and pRGN$_{SHV-18}$ with mmD247/248. PCR products were digested with SpeI and XmaI, ligated and transformed into *E. coli* S17-1 λpir

to create the donor cells used in matings.

Minimum Inhibitory Concentration (MIC) Determination

MICs were determined by broth microdilution using LB broth according to the CLSI guidelines [190].

Transformation Assays

Overnight cultures were diluted 1:100 in fresh LB and grown to an optical density (OD_{600}) of approximately 0.3–0.5. Following 15 minutes of incubation on ice, cultured cells were centrifuged at 3200 x g, and pellets were resuspended in one tenth volume of TSS buffer (LB, 10% polyethylene glycol, 5% dimethyl sulfoxide, 50mM Mg^{2+} at pH 6.5)[191]. A $100\mu L$ aliquot of cells was incubated with 10ng of RGN plasmid DNA. Plasmids were purified from the DH5αPro cloning host using a Qiagen QIAprep Spin Miniprep Kit and the concentration was determined using a Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen). Following 30 minutes of incubation on ice, cells were heat shocked at 42°C for 30 seconds, returned to ice for 2 minutes and recovered for 1.5h at 37°C in 1ml of SOC broth (HiMedia). For the chromosomal target assay, serial dilutions of cells were plated on LB+Cm to select for transformants. Plates were incubated overnight at 37°C, and the number of colony forming units (CFU) were enumerated the following day. Transformation efficiency was used to assess whether the given RGN plasmid was toxic to cells and was calculated as the CFU/mL per μg of DNA transformed.

For the episomal target assay, following recovery, cultures were washed in fresh LB, diluted 1:100 in LB supplemented with chloramphenicol to select for transformants and incubated for 16h at 37°C. Samples were washed in sterile PBS, serially diluted and plated on LB+Cm and LB+Cm+Cb or analyzed by flow cytometry. Colonies were enumerated the following day and plasmid loss was inferred by calculating the ratio of Cb^R+Cm^R CFUs to Cm^R CFUs.

Overnight cultures of RGN transformants were also diluted 1:100 in sterile PBS,

aliquoted in duplicates in a 96-well plate and immediately assayed using a BD LSR-Fortessa cell analyzer. Cells were consistently gated by side scatter and forward scatter across independent biological replicates. Fluorescence measurements were performed using a 488nm argon excitation laser. The GFP gate and laser voltages were initially determined using untreated pZE-*bla*$_Z$-*gfp* and EMG2 cells as positive and negative controls, respectively, and implemented across biological replicates. BD FACSDIVA software was used for data acquisition and analysis.

Sequence Analysis

Escape mutants from transformation assays were re-isolated by passaging surviving colonies onto LB+Cm+Cb. DNA isolation for escape sequencing analysis was performed by either extracting plasmid DNA from isolated escape mutants using the Qiagen QIAprep Spin Miniprep Kit or by amplifying the integrated target locus using primers mmD9/234 or mmD3/4 for *bla*$_{NDM-1}$ and *bla*$_{SHV-18}$, respectively. Sequencing was performed by Genewiz using the primers mmD112–115/153 and rcD11 for analysis of the RGN plasmids and mmD3 or mmD234 for examination of the integrated resistance genes.

Phagemid Purification

Phagemids encoding the RGNs were purified using the Qiagen QIAprep Spin Miniprep Kit (Qiagen) and transformed into *E. coli* DH5αPro along with the m13cp helper plasmid for generation of phagemid-loaded M13 particles [192]. Strains were inoculated and grown overnight in 250mL LB+Cm+Km to maintain m13cp and the phagemid, respectively. Cells were pelleted and the supernatant fluid containing the phagemid particles was passed through a 0.2μm filter. For all purifications except the ΦRGN*gyrA*$_{D87G}$ purification for the dose response curve, M13 phagemid particles were precipitated by the addition of 5% polyethylene glycol (PEG-6000) and 0.5M NaCl and incubation overnight at 4°C [193] and pelleted at 12,000 x g for 1h. Purified phagemid pellets were resuspended gently in 1/100th volume of SM buffer

(50mM Tris-HCl [pH 7.5], 100mM NaCl, 10mM $MgSO_4$) and stored at 4°C. For the $\Phi RGNgyrA_{D87G}$ purification for the dose response curve (Fig. 2b), M13 phagemid particles were precipitated [194] through the addition of concentrated HCl to pH 4.2 and subsequently pelleted at 13,000 x g for 15 minutes. The phagemid pellet was resuspended in 1/100th volume of water and concentrated NaOH was added to pH 7.0 to solubilize phagemid particles. Tris-HCl [pH 7.5], NaCl and $MgSO_4$ were added to reconstitute the composition of SM buffer.

Titers were measured by incubating sample dilutions with *E. coli* EMG2 for 30 minutes and enumerating transductants by plating on LB and LB+Km. Titers were defined in TFU_{100}/mL, which is the concentration of phagemid at which ~100% of a recipient population of an equivalent cell concentration would be transduced.

Phagemid Kill Assays

Cultures were inoculated and grown overnight in LB with appropriate antibiotics at 37°C with shaking. The following day, overnights were subcultured 1:100 into 3mL LB (no antibiotics) and grown at 37°C with shaking until the OD_{600} reached approximately 0.8. Cultures were diluted into LB to 10^8 CFU/mL for pNDM-1 and pSHV-18 assays (Fig. 1b, 2a) or 10^6 CFU/mL for $gyrA_{D87G}$ (Fig. 1c) and EHEC assays (Fig. 3a) and 245μL of the suspension was added to 5μL of purified phagemid stock in a 96-well plate and incubated static at 37°C. The number of viable cells in samples at each interval during the time-course or at 2h for endpoint assays was determined by serial dilution and spot plating onto LB, LB+Cb, and LB+Km to analyze cytotoxicity, plasmid loss, and phagemid delivery, respectively. Initial suspensions were also diluted and plated onto LB to quantify the initial bacterial inocula. Colonies were enumerated after 8–9h incubation at 37°C to calculate cell viability (CFU/mL) and averaged over three independent experiments. Non-linear curve fitting of the time-course to an exponential decay curve was performed using GraphPad Prism.

Galleria mellonella Model

Larvae of the model organism *Galleria mellonella*[168] were purchased from Vanderhorst Wholesale, Inc. (St. Marys, OH, USA) and received in the final larval instar for survival assays. Larvae were removed from food source, allowed to acclimate for at least 24h at room temperature in the dark, and used within 4 days of receipt. For all injections, a KDS100 (KD Scientific) or Pump 11 Elite (Harvard Apparatus) automated syringe pump was set to dispense a $10\mu L$ volume at a flow rate of $\sim 1\mu L/s$ through a 1mL syringe (BD) and 26G needle (BD). To prepare bacteria for injection, an overnight culture of E. coli O157:H7 43888 F' was subcultured in Dulbecco's Modified Eagle Medium (Gibco) for 4 hours at 37°C with shaking until OD_{600} ~ 0.6. Cultures were washed twice in PBS and diluted to a concentration of approximately 4×10^5 CFU/mL. In accordance with other studies [195], twenty larvae per treatment group were randomly selected based on size (150–250mg) and excluded based on poor health as evidenced by limited activity, dark coloration, or reduced turgor prior to experiments. Larvae were delivered injections without blinding of either PBS or bacteria behind the final left proleg. Approximately an hour after the first injection, SM buffer, antibiotic or ΦRGN treatment was administered behind the final right proleg. Larvae were incubated at 37°C and survival was monitored at 12h intervals for 72h, with death indicated by lack of movement and unresponsiveness to touch [168]. Kaplan-Meier survival curves were generated and analyzed with the log-rank test using GraphPad Prism.

LexA Reporter Assay

Overnight cultures of EMG2 WT, EMG2 pNDM-1 and EMG2 $gyrA_{D87G}$ containing the SOS-responsive reporter plasmid pZA3LG [164] were diluted 1:50 in LB and incubated with either SM buffer, ΦRGN$_{NDM-1}$ or ΦRGN$_{gyrA_{D87G}}$ at MOI ~ 5 for 2h at 37°C. Cultures were diluted 1:5 in $250\mu L$ of sterile PBS and analyzed using a BD LSR Fortessa cell analyzer, as above. BD FACSDIVA software was used for data acquisition and analysis was performed using FlowJo software.

Bacterial Matings

Donor and recipient strains grown overnight in LB with appropriate antibiotics were diluted 1:100 in fresh media and grown to approximately $OD_{600}=1$. Cells were pelleted, resuspended in sterile PBS and mating pairs were mixed at a donor to recipient ratio of 340±66:1. Mating mixtures were pelleted, resuspended in 20μL of PBS and spotted onto nitrocellulose filters placed on LB agar plates. Initial bacterial suspensions were serially diluted and plated on LB agar plates to quantify the initial inocula. Matings proceeded at 37°C for 3h with a single mixing step. At 90 minutes, mating mixtures were collected by vigorously vortexing the filters in 1mL sterile PBS. Cells were pelleted, resuspended in 20μL PBS and re-seeded onto filters and incubated as above for the remaining 90 minutes. At the end of the 3h mating, cells were again recovered by vigorously vortexing the filters in 1mL sterile PBS. Mating mixtures were serially diluted in PBS and plated onto LB+Cb to select for total number of Cb-resistant recipient cells and LB+Cb+Cm to select for transconjugants. Colonies were enumerated following overnight incubation at 37°C to determine viable cell counts and were averaged over nine independent biological replicates.

Synthetic Consortia Remodeling

E. coli CJ236, EMG2 pNDM-1 and RFS289 strains grown overnight in LB with appropriate antibiotics were diluted 1:100 into fresh LB (no antibiotics) and grown to OD_{600} ~0.8. Cultures were seeded into fresh LB such that the initial mixture contained ~1x10^6 CFU/mL of each strain and 245μL of the suspension was added to 5μL of SM buffer or purified ΦRGN$_{NDM-1}$ or ΦRGN$_{\textit{gyrA}_{D87G}}$ in triplicate in a 96-well plate and spotted onto LB +Cm, +Sm and +Ofx to quantify the initial concentration of CJ236, EMG2 pNDM-1 and RFS289, respectively. Samples were then incubated, plated and enumerated as in phagemid kill assays. The composition of the synthetic ecosystem under each treatment condition was determined by counting viable colonies on plates selective for each strain as above and data were calculated as viable cell concentration (CFU/mL) averaged over three biological replicates (Fig.

4).

Data Analysis and Statistics

All data were analyzed using GraphPad Prism version 6.0 (GraphPad Software, San Diego, CA, USA, www.graphpad.com).

Chapter 3

Domesticating Commensal *Bacteroides*

The following chapter is adapted from the 2015 Cell Systems publication to which Alex C. Tucker and I contributed equally as first authors. [196]

Abstract

Engineering commensal organisms for challenging applications, such as modulating the gut ecosystem, is hampered by the lack of genetic parts. Here, we describe promoters, ribosome-binding sites, and inducible systems for use in the commensal bacterium *Bacteroides thetaiotaomicron*, a prevalent and stable resident of the human gut. We achieve up to 10,000-fold range in constitutive gene expression and 100-fold regulation of gene expression with inducible promoters and use these parts to record DNA-encoded memory in the genome. We use CRISPR interference (CRISPRi) for regulated knockdown of recombinant and endogenous gene expression to alter the metabolic capacity of *B. thetaiotaomicron* and its resistance to antimicrobial peptides. Finally, we show that inducible CRISPRi and recombinase systems can function in *B. thetaiotaomicron* colonizing the mouse gut. These results provide a blueprint for engineering new chassis and a resource to engineer *Bacteroides* for surveillance of or therapeutic delivery to the gut microbiome.

3.1 Introduction

Bacteroides species are prominent Gram-negative anaerobic symbionts of the mammalian gut microbiome [1], comprising 30% of culturable anaerobes in the human gastrointestinal tract [197]. Of the Bacteroides genus, Bacteroides thetaiotaomicron is both prevalent (present in 46% of humans) [1] and abundant (up to 10^{10} per g stool) [197], making it a promising organism for both understanding and manipulating the gut environment. Stable and robust colonization of the densely populated gut environment is facilitated by the metabolic diversity of *Bacteroides* [198]. Specifically, *B. thetaiotaomicron* and its relatives are equipped with an extensive repertoire of saccharolytic enzymes and serve as primary fermenters of host-, diet- or microbially-derived polysaccharides [199]. Moreover, members of *Bacteroidetes*, the phylum to which *B. thetaiotaomicron* belongs, were among the most stable components of the human gut microbiota in a 5-year longitudinal study [6], making them useful candidates for long-term cellular diagnostics and therapeutics in the gastrointestinal tract. To date, multiple microorganisms have served as chassis for engineered microbial therapies of human disease. Recombinant, attenuated strains of *Salmonella* spp. [200] and *Listeria monocytogenes* [201] were used as vectors for anti-cancer therapies in several human trials. *Lactococcus lactis* has been implemented as a chassis for the production of therapeutic molecules targeting human inflammatory diseases [202]. In mouse models, strains of *Escherichia coli* were engineered to produce molecules that reduce food intake and obesity [104]. However, compared to *E. coli* [109] and *L. lactis* [202], which undergo depletion or clearance within days of administration, *Bacteroides* populations exhibit low variation in abundance and long-term colonization [198].

Genetic parts and circuits enable control over the level and timing of expression of multi-gene traits in response to environmental conditions. Recently, new techniques in DNA construction and high-throughput screening have led to a resurgence in part design, including a revisiting of the organization of the "expression cassette" [203, 204]. For model organisms, large libraries and computational models of promoters, ribosome binding sites, and terminators enable fine control of multi-gene systems

[204]. Insulators have been designed for integration between parts to ensure that parts can be swapped without impacting function [205, 206]. However, a challenge with engineering non-model organisms has been a lack of these tools, which limits the sophistication of the systems that can be constructed.

Few genetic parts and inducible systems are available for *B. thetaiotaomicron* and its relatives. Previous efforts have co-opted natural glycan-sensing systems [207] or the classic *E. coli lac* operon [208] for inducible genetic systems, yielding systems that span a 10- to 50-fold range of expression. Replicative plasmids [209] and integrative transposons [210] have been built for the introduction of heterologous genes. However, unlike most other prokaryotes, the unique major sigma factor in *Bacteroides* binds to a -33/-7 consensus sequence (TTTG/TAnnTTTG) [211, 212]. Moreover, the strength of translation initiation is poorly correlated with the level of ribosome binding site (RBS) complementarity to the 16S rRNA of the host organism [213]. Compared to the *E. coli* RBS, *Bacteroides* RBS strength is more sensitive to secondary structures [214], depleted in GC content [213], and predicted to rely more heavily on interactions with ribosomal protein S1. These unique promoter and RBS architectures in *Bacteroides* preclude the direct incorporation of genetic systems developed in other organisms [209]. A lack of genetic part libraries hinders the introduction of multi-gene pathways, such as those that could produce a metabolic product designed to treat disease. Multiple cellular states are naturally maintained within *Bacteroides* populations via reversible recombinases that vary expression of cell surface polysaccharides [215]. Recombinases have been harnessed to build counters and integrated memory-and-logic devices [216–219]. By connecting these switches to environmental sensors, cellular memory can be used to infer exposure to a particular set of conditions. This is useful if direct readout of a reporter from a microenvironment is not possible. For example, an engineered toggle switch enabled *E. coli* to record exposure to antibiotics in the mouse gut [109]. In addition to toggle switches, memory has recently been achieved using dynamic genome editing [220].

CRISPR-Cas9 technologies have revolutionized genome editing because their ease of reprogramming to target user-defined DNA sequences with a single guide RNA

(sgRNA) [221]. CRISPR interference (CRISPRi) employs a catalytically inactive version of the endonuclease Cas9 (dCas9) to regulate gene expression at target DNA sequences by blocking transcription by RNA polymerase [222]. The specificity of dCas9 repression is governed by sequence homology and is independent from host machinery [221]. This has enabled implementation of CRISPRi in both bacterial and eukaryotic systems [222], permitting both the construction of synthetic gene circuits as well as the study of natural biological networks [223].

Here, we expand the set of genetic tools necessary to precisely and robustly engineer *B. thetaiotaomicron* for microbiome applications. We report a library of biological parts, comprised of constitutive promoters, inducible promoters, and RBSs that each span output-dynamic ranges of several orders of magnitude (Fig. 3-1). Constitutive promoters and RBSs were used to characterize the input expression levels required to generate recombinase-based DNA-encoded memory in *B. thetaiotaomicron*. Externally switchable DNA-based memory devices were then constructed by integrating inducible promoters with recombinases. Additionally, inducible promoters were used to control CRISPRi-based regulation of synthetic and endogenous genes. Finally, we validated the function of these tools in bacteria that have colonized the gut of mice. This includes circuits that respond to stimuli *in vivo* and either record exposure by altering gene expression or permanently modifying the genome. These devices demonstrate that *B. thetaiotaomicron* can be used as a platform for predictable gene expression and circuit design for microbiome engineering.

3.2 Landing Pads for Genetic Part and Device Characterization

The integration vector pNBU2 was employed for most of the genetic parts in this chapter to ensure genetic stability of the constructs (Fig. 3-2). The pNBU2 plasmid encodes the *intN2* tyrosine integrase, which mediates sequence-specific recombination between the *attN* site of pNBU2 and one of two *attBT* sites located in the 3' ends of
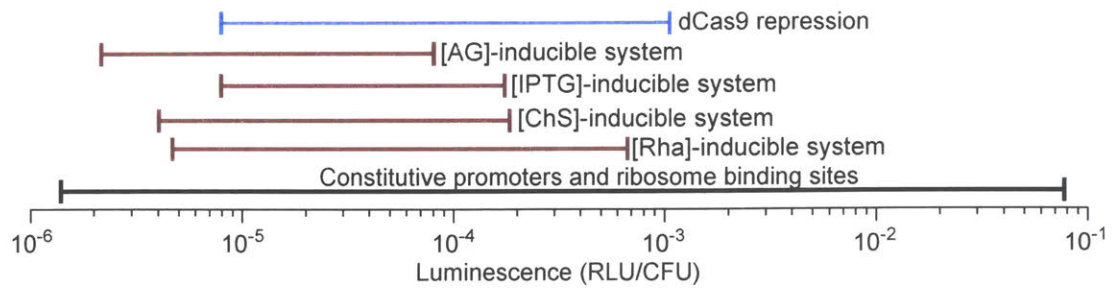
Figure 3-1: ***Bacteroides* Genetic Part Summary:** The ranges of gene expression are shown for the different gene regulation systems described in this chapter. AG, arabinogalactan; CS, chondroitin sulfate; IPTG, isopropyl beta-D-1-thiogalactopyranoside; Rha, rhamnose; dCas9, catalytically inactive Cas9 endonuclease

the two tRNA^{Ser} genes, BT_t70 and BT_t71, on the *B. thetaiotaomicron* chromosome [210]. Insertion of the pNBU2 plasmid inactivates the tRNA^{Ser} gene, and simultaneous insertion into both BT_t70 and BT_t71 is unlikely due to the essentiality of tRNA^{Ser}.
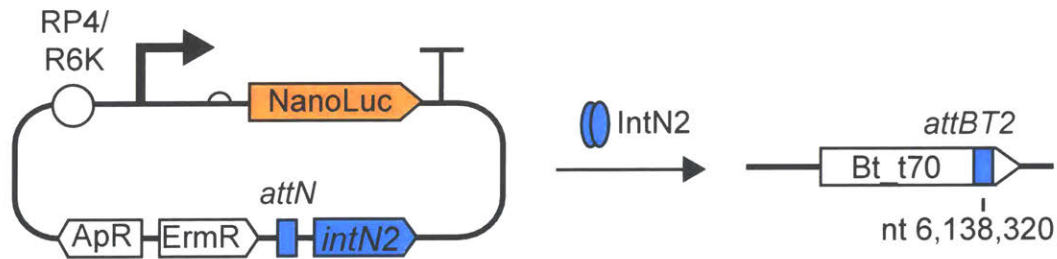
Figure 3-2: :

pNBU2: Tyrosine integrase IntN2 catalyzes stable integration of pNBU2-based expression constructs into one of two *attBT2* sites in the *B. thetaiotaomicron* genome [210]. The two *attBT2* sites (*attBT2-1* at nt 6,217,227 and *attBT2-2* at nt 6,138,320) are in the 3' ends of tRNA^{Ser} genes (BT_t71 and BT_t70, respectively). ApR, ampicillin resistance cassette; ErmR, erythromycin resistance cassette; RP4, origin of transfer; R6K, origin of replication; NanoLuc, luciferase

3.2.1 Development of a Broad Host Range Mobilization Vector in *Bacteroides*

The pNBU2 expression vector is effective for transfer of genetic constructs into *B. thetaiotaomicron*. However, integration efficiency in some other *Bacteroides* spp. is

reduced, especially in *B. fragilis* where no transconjugants could be obtained (Fig. 3-3A). After construction of multiple engineered *B. thetaiotaomicron* strains, we found that pNBU2 occasionally catalyzes integration into a third site (*attBT2-3*), leading to disruption of the BT4401 open reading frame (Fig. 3-3B). In light of these suboptimal characteristics, we developed an alternative mobilization vector based on the NBU1 conjugative transposon. Discovered prior to NBU2, the NBU1 transposon can be mobilized into *Bacteroides* species by conjugative machinery *in trans* [224]. The IntN1 integrase catalyzes site specific integration of transposon-encoded *attN1* into the 3' end of tRNA^{Leu} (*attBT1*) [225]. Moreover, characterization of IntN1 revealed that a C to G mutation within *attN1* can increase site-specific integration into *attBT*1, while reducing non-specific integration events [226]. The integrase and modified *attN* sites of NBU1 were exchanged for those from NBU2 to yield the pNBU1 integration vector. Conjugation efficiencies of pNBU1 varied depending on the recipient species; efficiency was increased in *B. thetaiotaomicron*, *B. fragilis* and *B. ovatus*, comparable in *B. uniformis* and *B. eggerthii*, and decreased in *B. vulgatus* compared to pNBU2 (Fig. 3-3A). Compared to pNBU2, pNBU1 did not produce any detectable off-site integration events in *B. thetaiotaomicron* (Fig. 3-3B). Finally, the IPTG-inducible system described in Fig. 3-7 was encoded onto pNBU1 and transferred into several *Bacteroides* spp. Luminescence in the resultant strains was IPTG-inducible and comparable across the species (Fig. 3-3C), suggesting that pNBU1 could be a broad applicable vector system in the *Bacteroides* genus.

3.3 Expression Control Through Promoter and RBS Design

To expand the range of constitutive gene expression that can be implemented in *Bacteroides*, we constructed and characterized promoter-RBS combinations using the NanoLuc luciferase as a reporter (Fig. 3-4A). Four promoter variants were constructed based on the constitutive promoter for the *B. thetaiotaomicron* housekeeping

A

C

B

Site Specificity		
Site	**NBU1**	**NBU2**
attB1	100%	0%
attB2-1	0%	30.5%
attB2-2	0%	63%
attB2-3	0%	6.5%
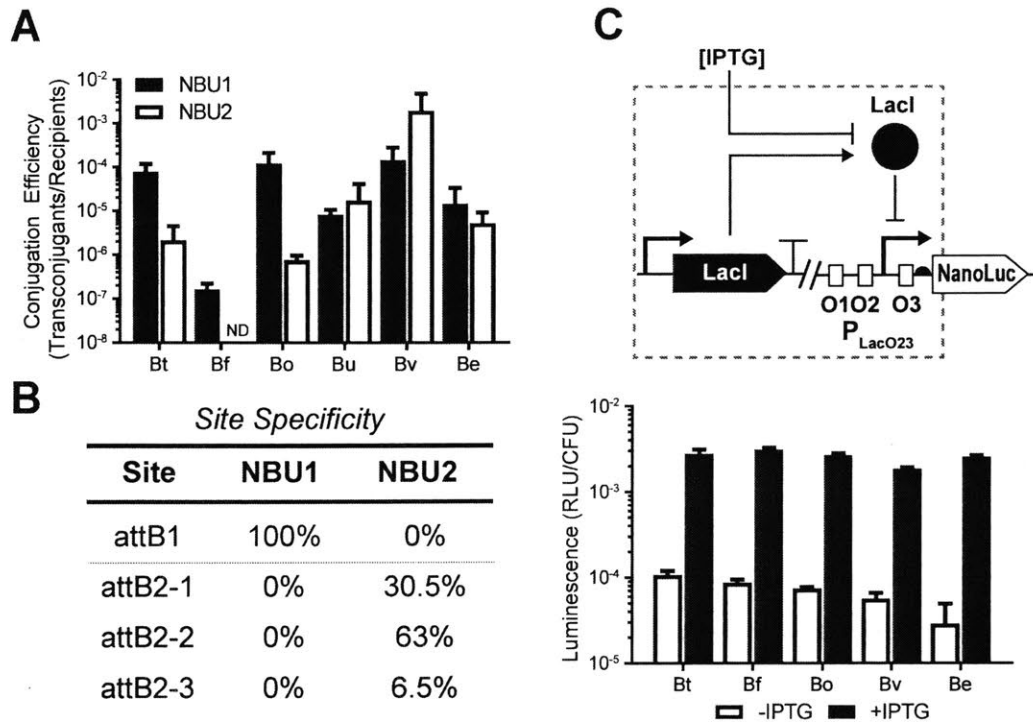
Figure 3-3: :

Broad-Host Range pNBU1 Vector System: (A) Conjugation efficiency of pNBU1 and pNBU2 were compared in B. thetaiotaomicron (Bt), B. fragilis (Bf), B. ovatus (Bo), B. uniformis (Bu), B. vulgatus (Bv), and B. eggerthii (Be). Conjugation efficiency was measured as the number of transconjugants following overnight mating divided by the total number of recipient cells. (B) Using integration site-specific primers for PCR, B. thetaiotaomicron transconjugants were screened to determine the site-specificity of each integrase. NBU1 catalyzed integration in a unique site at the 3' end of a leucine tRNA gene. NBU2 integrated favoured integration in serine tRNA genes (attB2-1/2), but some constructs were inserted in BT4401 (attB2-3). (C) The IPTG-inducible expression system described in Fig. 3-7 was encoded onto pNBU1 and transferred into multiple Bacteroides spp. Luminescence activity was measured in the presence and absence of IPTG. The genetic circuit performed efficiently in all species.

sigma factor BT1311 (P_{BT1311}) [212]. Specifically, a 26-bp sequence was substituted or inserted into P_{BT1311} in regions composing and surrounding the -33 and -7 promoter sequences (Experimental Details, Fig. 3-4B), regions known to be important for *B. thetaiotaomicron* promoter activity [211]. With NanoLuc as a reporter, the P_{AM} promoters spanned a 20-fold range of expression and had decreased expression levels relative to the P_{BT1311} parent promoter. For comparison to prior work, we also

measured the activities of the previously reported promoter-RBS pairs, P_{cfxA}, P_{cfiA}, P_1 and P_{cepA} [213, 227–229] (Fig. 3-4C).

The P_{AM} promoters were then combined with RBSs of varying strength to increase the range of expression levels. Previously characterized RBSs GH022, GH023, and GH078 [213] have an expression range of expression less than one order of magnitude (Fig. 3-4C). We selected a ribosomal protein RBS (rpiL*) and constructed a weak RBS (RC500) to increase the range of our available RBSs (Fig. 3-4C, Experimental Details). This RBS library spanned a $>10^2$-fold range when paired with each P_{AM}-derived promoter. When combined, these P_{AM} promoters and RBSs could achieve expression levels over a 10^4-fold range.

To identify a set of RBSs for fine-tuning gene expression in *B. thetaiotaomicron*, we generated three randomized rpiL* RBS libraries targeting the most conserved positions of the *Bacteroides* ribosomal protein RBSs [213]. For each library, we targeted 3 nucleotides in and around the rpiL* RBS Shine Delgarno sequence. These positions are within or near the RBS region predicted to interact with the ribosomal S1 protein (nt -21 to -11 relative to the start codon of NanoLuc, Figure 1C) [230]. We achieved 67-80% coverage of the 64 potential members in each library, resulting in 142 RBS sequences (Figure 1E, Table S1). These RBSs were screened and sequenced and a set of 8 was identified that span 10^3-fold expression range in approximately even increments (Table S2).

We observed only a weak positive correlation between the minimum free energy of RBS folding (Experimental Details) and expression of the NanoLuc reporter ($r^2 = 0.19$) in the rpiL* library (Figure 1F). To visualize the impact of GC content on RBS strength within this library, we generated frequency logos for each nucleotide targeted in the library. The strongest RBSs were GC-depleted relative to the overall library and the weakest RBSs (Figure 1G). Our RBS libraries highlight the distinct AT enrichment of strong *Bacteroides* RBSs compared to other bacterial species, which results in part failure when constructs are transferred into *Bacteroides* from other species [213].
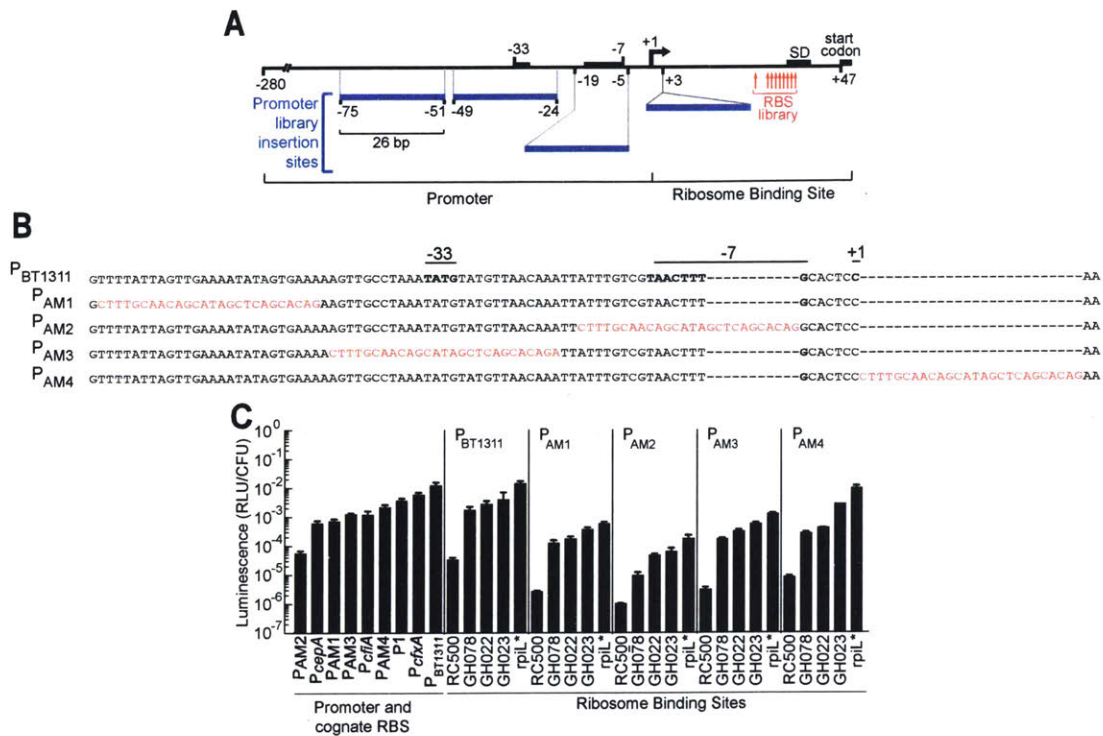
Figure 3-4: **Constitutive promoters and ribosome binding sites for the construction of gene expression libraries:** (A) The putative -33 and -7 regions of the P_{BT1311} promoter, the Shine-Dalgarno sequence, and the start codon are indicated by black boxes. Numbers below the black boxes represent nucleotide locations relative the P_{BT1311} transcription start site. The 26 bp sequence introduced in the P_{AM} promoters is shown as blue boxes (See also Figure S1). Numbers at the edges of the blue boxes indicate the P_{BT1311} nucleotides replaced or the insertion site within the promoter. The location of residues randomized in the rpiL* RBS library are indicated with red arrows (for library A: nt -14, -13, -12; for library B: nt -21, -18, -15; and for library C: nt -17, -16, -11; nt numbering is relative to the translation start site). (B) Promoters P_{AM1}, P_{AM2}, P_{AM3}, and P_{AM4} were constructed by introducing a 26 bp sequence (red) at 4 locations in the constitutive BT1311 promoter (P_{BT1311}). Predicted -33, -7, and +1 sites of the P_{BT1311} promoter are shown in bold. (C) Activity was measured for a set of constitutive promoters and their cognate RBSs. For P_{AM1}, P_{AM2}, P_{AM3}, P_{AM4}, the BT1311 RBS was used. Furthermore, P_{BT1311}, P_{AM1}, P_{AM2}, P_{AM3}, and P_{AM4} were combined with RBSs of varying strengths. Gene expression was measured using a luciferase reporter (NanoLuc) and reported as relative light units / colony forming unit (RLU/CFU).

3.4 Genetic Sensors and Inducible Systems

To create inducible systems for use in *B. thetaiotaomicron*, we adapted parts from the large repertoire of systems that govern carbohydrate utilization [199]. In *B.*
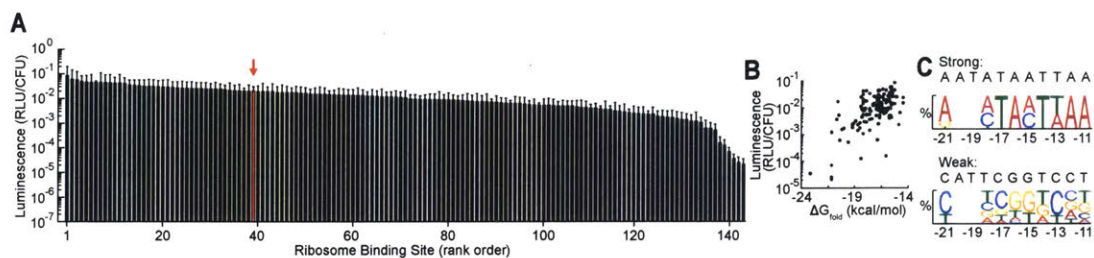
Figure 3-5: **Characterization of Random RBS Library:** (A) Three large RBS libraries were constructed and combined with promoter P_{BT1311}. For reference, the parent rpiL* RBS is indicated with a red arrow. (B) The strength of each RBS was compared to the predicted free energy of folding for the mRNA (ΔG_{fold}). (C) A consensus strong RBS and weak RBS were generated for the rpiL* RBS library using frequency logos that included the 11 strongest and 11 weakest RBSs (residue locations are stated relative to the translation start site). Frequency logos were constructed by comparing the frequency of each nucleotide at each position in that group with the frequency of that nucleotide in that position in the full library. Position -20 and -19 were not randomized and are not shown in the frequency logos.

thetaiotaomicron, rhamnose metabolism is mediated by the transcriptional activator RhaR, which activates transcription at the P_{BT3763} promoter [231]. To assay the functionality of P_{BT3763} as an inducible system, we cloned 250bp of the promoter-RBS region upstream of the start codon of BT3763 into the pNBU2 expression vector to drive expression of NanoLuc. Gene expression was conditional on the concentration of rhamnose and demonstrated a response curve with an output dynamic range of 10^4-fold (Fig. 3-6A). Fitting the response curve to a Hill function revealed a threshold K of 0.3 mM and a Hill coefficient n = 1.4.

Two-component systems are signal-transduction mechanisms widespread in bacteria for sensing external stimuli. *Bacteroides* sp. possess a unique variant of these systems, called hybrid two-component systems, that incorporate both the sensor histidine kinase and response regulator of classical two-component systems into a single polypeptide chain [232]. Putative hybrid two-component systems, BT3334 and BT0267, were identified in transcriptomic studies to control expression of the chondroitin sulfate (ChS)-inducible P_{BT3324} promoter and arabinogalactan (AG)-inducible P_{BT0268} promoter, respectively [199, 233]. Chondroitin sulfate induction of P_{BT3324} and arabinogalactan induction of P_{BT0268} led to a 60-fold and 29-fold regulation of
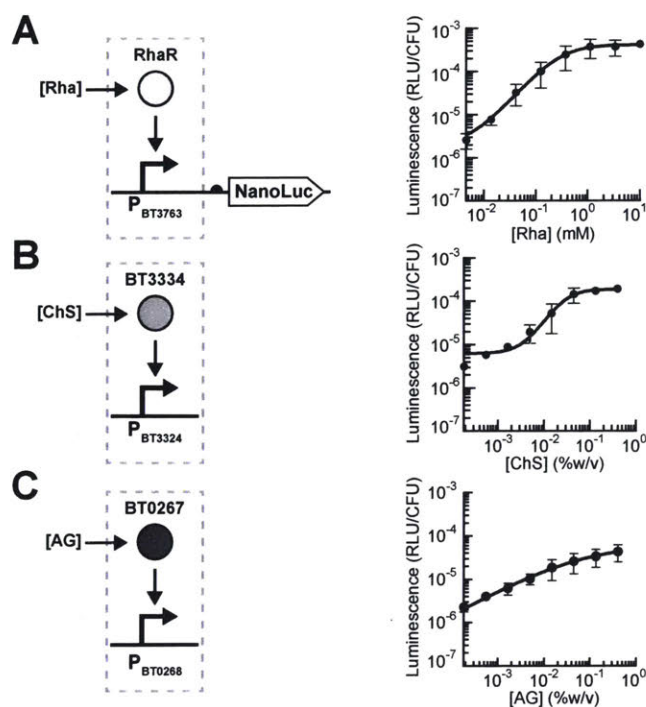
Figure 3-6: **Design and characterization of genetic sensors.** (A-C) Response curves for NanoLuc under the regulated control of the rhamnose- (Rha) (A), chondroitin sulfate- (ChS) (B), or arabinogalactan- (AG) (C)) inducible promoters. Inducer concentrations were applied as follows: three-fold serial dilutions starting at 10mM Rha (A); three-fold serial dilutions starting at 0.4% for ChS (B), and AG (C). The leftmost data point in each plot represents the background luminescence in the absence of inducer. Response curves were fit to a Hill function (solid lines).

output gene expression, respectively (Fig. 3-6B and C).

Next, we developed an IPTG-inducible system based on the *E. coli* LacI system. Our design expands upon on a previously developed IPTG-inducible system in *Bacteroides* [208] by investigating the position effects of operator sites on gene expression. Pairs of LacO1 operator sites were inserted in the strong P_{cfxA} promoter in three locations (Fig. 3-7A), and the LacI repressor was expressed from the BT1311 promoter. Compared to the unmodified P_{cfxA} promoter, the addition of synthetic operator sites diminished the maximum expression of NanoLuc (Fig. 3-7B). This strategy produced two IPTG-inducible promoters that with thresholds at $K = 86\mu M$ (P_{LacO13}) and $K = 6\mu M$ (P_{LacO23}). The induction of these systems elicited an 8- and 22-fold change in gene expression, respectively.
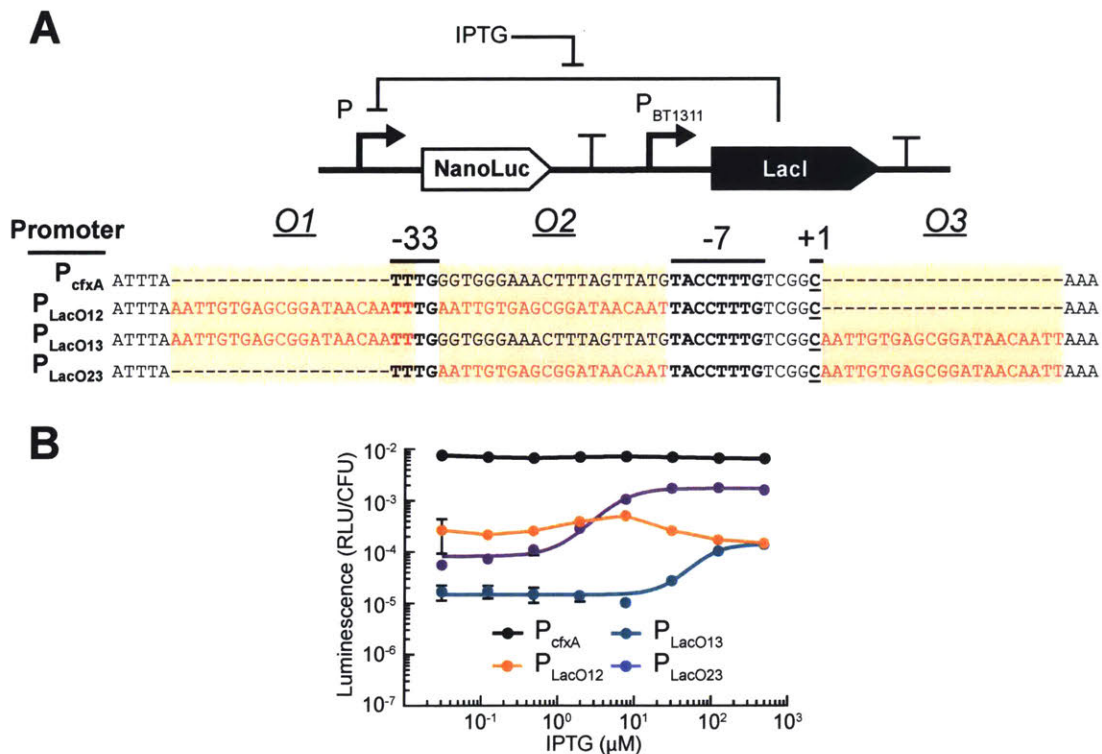
Figure 3-7: **Development of an IPTG-inducible promoter system.** (A) Synthetic IPTG-inducible promoters were constructed by placing LacO1 operator sites (red) upstream of the -33 element (O1), between the -33 and -7 elements (O2) and/or directly downstream of the transcription start site (O3) of the strong P_{cfxA} promoter. Predicted -33, -7 and +1 sites are shown in bold. These promoters are regulated by the E. coli LacI repressor expressed from P_{BT1311}. (B) Response curves for the synthetic IPTG-inducible systems. Cells were incubated with no inducer or four-fold serial dilutions of IPTG starting at 500 μM. Data sets for P_{LacO13} and P_{LacO23} were fit to a Hill function (solid line). Error bars represent the standard deviation of three biological replicates made on three different days (n=3).

As the orthogonality of genetic parts is crucial for their simultaneous use, we tested the degree of cross-talk between each inducible system by incubating each engineered strain with the full set of carbohydrate inducers. The inducers themselves bear little structural similarity: rhamnose, a methyl-pentose sugar; ChS, a sulfated glycocosaminoglycan composed of chains of acetylgalactosamine and glucuronic acid residues; AG, a polysaccharide composed of arabinose and galactose units; and IPTG, a molecular mimic of allolactose. Functionally, each inducible system was highly orthogonal to each other, with no cross-reactivity observed with any of the combinations

(Fig. 3-8).
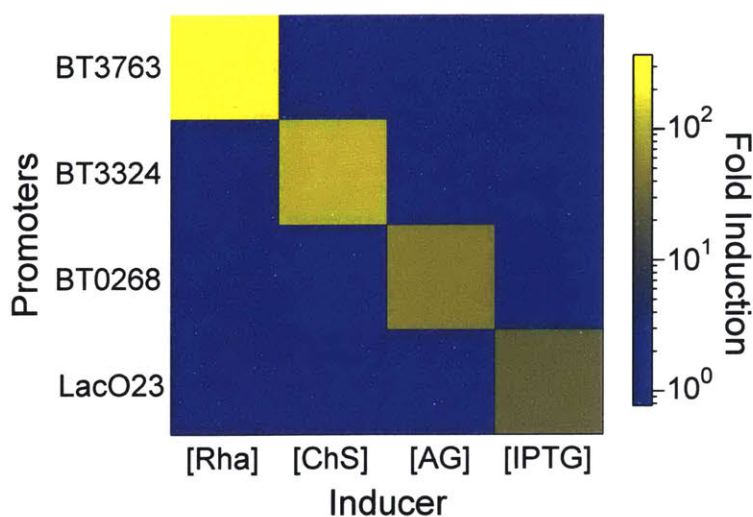
Figure 3-8: **Orthogonality matrix of *Bacteroides* inducible promoters.** Orthogonality matrix of sugar-inducible genetic systems incubated with 10mM rhamnose (Rha), 0.2% chondroitin sulfate (ChS), 0.2% arabinogalactan (AG), or 100μM IPTG compared to no inducer. Error bars represent the standard deviation of three biological replicates made on different days.

3.5 Synthetic Genetic Memory

To equip *B. thetaiotaomicron* with permanent genetic memory, we used serine integrases, which catalyze unidirectional inversion of the DNA sequence between two recognition sequences (Fig. 3-9A) [234]. Recently, 11 orthogonal integrases and their recognition sequences were characterized in *E. coli* [219], and a DNA "memory array" composed of a linear concatenation of integrase recognition sequences was used to record the expression of one or multiple integrases in response to a stimulus.

We identified serine integrases that function in *B. thetaiotaomicron* by cloning the integrases into a strong constitutive expression vectors (P_{AM4}-rpiL*, 1.2 x 10^{-2} RLU/CFU). To provide a stable, single-copy record of DNA inversion, we incorporated the DNA memory array containing the integrase recognition sequences into the *B. thetaiotaomicron* chromosome (Fig. 3-9, Experimental Details). Integrase expression vectors were conjugated into the *B. thetaiotaomicron* memory array strain.

Genomic DNA was isolated from transconjugants and analyzed by PCR to detect flipping. Four integrases, Int7, Int8, Int9, and Int12, each catalyzed recombination at the respective recognition sequence in the memory array (Fig. 3-9D), and DNA inversion was not detected in the absence of an integrase.

To create an inducible memory switch, we cloned Int12 under the control of the rhamnose-inducible promoter with the rpiL*RBS C51 (Fig. 3-9 E, Fig. 3-5A). The Int12 recombinase switch responded to increasing concentrations of rhamnose (Fig. 3-9F) within 2 hours (Fig. 3-9G), with no background detected in the absence of inducer. Notably, expression of Int12 did not impact growth of *B. thetaiotaomicron*.

A

attB   attP

Int12

182 nt

B

Integrase

attBT2-2   attBT2-1
6,138,320   6,217,227

B. thetaiotaomicron
chromosome

Memory Array
2,660,694

C integrase target sites

2 3 4 5 7 8 9 10 11 12 13

BT2113   Memory Array

BT2113   BT2114 BT2116
          BT2115 BT2117

D   Int7 Int8 Int9 Int12 bp

+Int   − 400
                                  − 200

         Int7 Int8 Int9 Int12 bp

−Int   − 400
                                  − 200

E

[Rha] → RhaR

P_BT3763

Int12

attB   attP

Int12

F

Recombination (Flipped/Total Amplicons)

1.0

0.5

0.0

10^{-2} 10^{-1} 10^{0} 10^{1}

Rhamnose (mM)

G

Recombination (Flipped/Total Amplicons)

1.0

0.5

0.0

0 2 4 6 8

Time (hr)

Figure 3-9: **Synthetic Genetic Memory.** A) Integrases mediate recombination of DNA between integrase binding sites ($attB/attP$), resulting in the inversion of the intervening spacers. (B) Schematic of the location of the promoter-RBS-integrase system and the memory array cassettes in the $B.$ $thetaiotaomicron$ chromosome. (C) Integrase-mediated DNA inversion at each integrase target sequence in the memory array cassette is detected by PCR. Primer pairs (arrows) anneal to the interface of the integrase recognition sites and to the spacer region between recognition sites. PCR amplification occurs only after an inversion event (solid lines below the primer arrows indicate expected amplicons). (D) Representative PCR products are shown after recombination with integrases Int7, Int8, Int9, or Int12. – indicates no integrase, + indicates the integrase is present. P_{AM4}-rpiL* was used to control expression of each integrase. (E) Schematic of the rhamnose-inducible recombinase circuit. Transcriptional activator RhaR, produced from the endogenous locus, is activated in the presence of rhamnose causing expression of Int12 from P_{BT3763}. Int12 mediates recombination between the Int12 $attB$ and $attP$ recognition sequences. (F) Response curve of Int12 memory circuit. Int12 was placed under the control of a subset of P_{BT3763}-rpiL*C51. Inducer concentrations were nine-fold serial dilutions starting at 10mM rhamnose. The leftmost data point represents the recombination in the absence of inducer. Cells were grown 8 hours at 37°C before harvesting cells and isolating DNA. Absolute quantities of flipped and unflipped memory array in genomic DNA were determined by qPCR using standard curves (Experimental Details). The recombination ratio is expressed as the ratio of cells containing a flipped memory array (Flipped) divided by the sum total of cells containing a flipped or unflipped array (Total). Data were fit with a Hill function. (G) Int12-mediated recombination versus time. Cells were induced with 10 mM rhamnose at t=0. Absolute quantities of flipped and unflipped memory array in genomic DNA were determined by qPCR using standard curves (Experimental Details). Recombination ratios were determined as in F. Data were fit with a sigmoidal dose-response function. For F-G, error bars represent the standard deviation of three biological replicates made on different days.

3.6 CRISPRi-Mediated Gene Knockdown

CRISPRi can provide a facile toolbox for constructing synthetic gene circuits and modulating endogenous genes in $B.$ $thetaiotaomicron$. To demonstrate the use of CRISPRi-mediated gene knockdown for synthetic constructs, production of dCas9 was regulated by the IPTG-inducible P_{LacO23} system. sgRNAs were constitutively expressed from the P_1 promoter and were designed to target the coding sequence of NanoLuc (NL1-4) or the P_{cfiA} promoter (PR1-2) (Fig. 3-10A). A nonsense sgRNA (NS) with no sequence identity to either P_{cfiA} or NanoLuc was used as a negative control. All specifically targeted sgRNAs repressed the expression of NanoLuc (Fig.

3-10B) by 20-45 fold with IPTG induction of dCas9 expression (Fig. 3-10C), thus implementing genetic NOT gates in *B. thetaiotaomicron*. The IPTG-to-NanoLuc response function of sgRNAs targeting the coding sequence or promoter exhibited similar Hill coefficients and lower dissociation constants to the IPTG-to-NanoLuc transfer function of the P_{LacO23} promoter on its own (n = 1.1 to 1.4; K = 0.6 to 1.4μM IPTG).
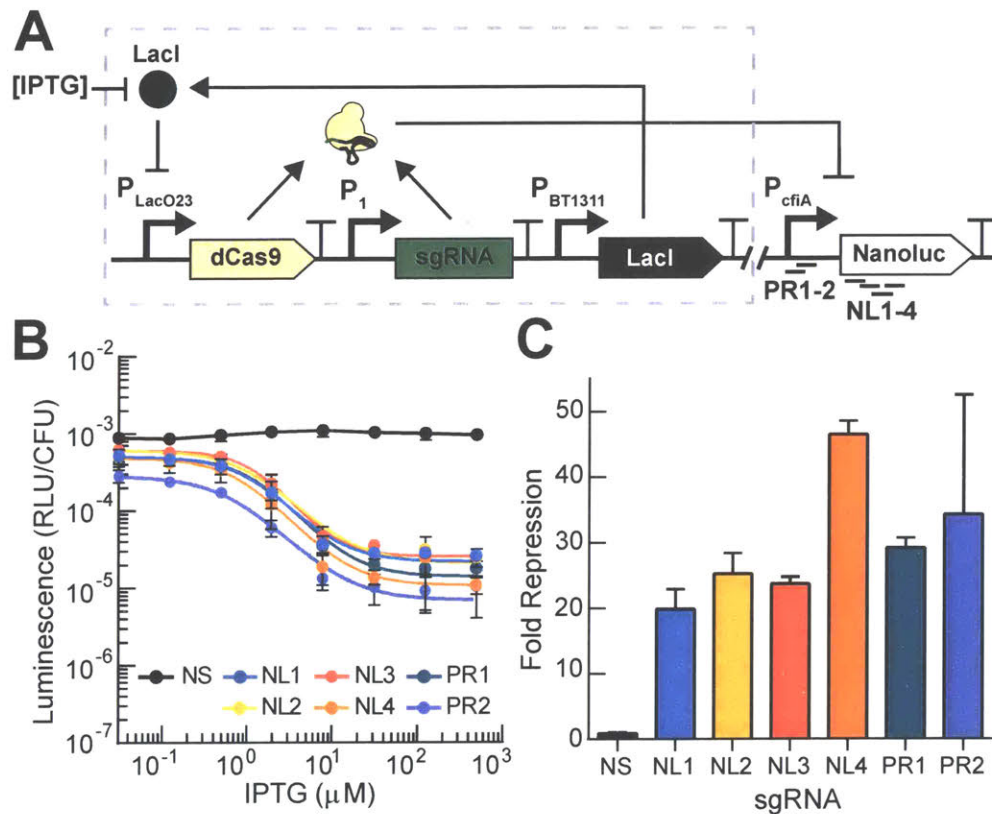
Figure 3-10: **CRISPRi-mediated repression of recombinant genes.** (A) Schematic of dCas9-based repression of NanoLuc. Addition of IPTG induces expression of dCas9, which complexes with constitutively expressed sgRNA targeting the coding sequence of NanoLuc (NL1-4) or the P_{cfiA} promoter (PR1-2). The plasmid backbone separates the NanoLuc cassette and the IPTG-inducible CRISPRi system. (B) Response curves of dCas9-mediated targeting the coding sequence of NanoLuc (NL1-4), the promoter (PR1-2) or a nonsense sequence (NS). Fourfold serial dilutions of IPTG starting at 500 μM or no inducer were added to cultures. Response curves were fit to a Hill Function (solid lines). (C) Fold repression elicited by various gRNAs in the presence (500 μM) of inducer. Bars are coloured to correspond to part B.

To demonstrate the programmable knockdown of endogenous genes in *B. thetaio-*

taomicron, we designed sgRNAs to target mechanisms implicated in the resilience of *Bacteroides* in the human microbiota. In *B. thetaiotaomicron*, LpxF, the gene product of BT1854, is required for resistance to inflammation-associated cationic antimicrobial peptides, such as polymyxin B [235]. Using the minimum inhibitory concentration (MIC) of polymyxin B as a phenotypic readout, we designed an sgRNA to specifically suppress BT1854 expression. In cells containing the sgRNA targeted against BT1854 (dCas9$_{BT1854}$), the induction of dCas9 with led to sensitization of the cells to polymyxin B treatment (8 to 16-fold decrease in MIC), while wild-type *B. thetaiotaomicron* and strains containing dCas9$_{NS}$ demonstrated high levels of polymyxin B resistance in the presence or absence of dCas9 induction with IPTG (Fig. 3-11D-E).

Next, we explored dCas9-mediated repression of carbohydrate-utilization pathways in *B. thetaiotaomicron*, which are important for the bacterium's ability to successfully and persistently colonize the mammalian gut. The hybrid two-component sensor BT1754 regulates the BT1757-1763/BT1765 fructose-containing polysaccharide utilization locus, and BT1754 is essential for growth on fructose as the sole carbon source [236]. We designed a specific sgRNA to repress BT1754 and integrated this system into the *B. thetaiotaomicron* genome along with an IPTG-inducible dCas9 cassette (dCas9$_{BT1754}$) (Fig. 3-11D,F). Induction of dCas9$_{BT1754}$ did not affect the growth rate of cells on MM-glucose compared to WT cells and dCas9$_{NS}$. The generation time $G = (\log_{10}2 \times t)/\log_{10}(B/B_o) \sim 1hr$ (where t is the time interval, and B_o and B are the initial and final concentrations of bacteria, respectively), indicating that neither dCas9 induction nor repression of BT1754 impacts growth on glucose media (Fig. 3-11F). However, induction of dCas9$_{BT1754}$ drastically decreased the growth rate of the cells in MM-fructose (G = 4.7 hr) while the growth of WT and dCas9$_{NS}$ cells in MM-fructose remained similar (G = 1 hr) to growth in MM-glucose (Fig. 3-11F). Thus, inducible dCas9-mediated repression of endogenous genes can alter both the resistance and metabolic profiles of *B. thetaiotaomicron*.
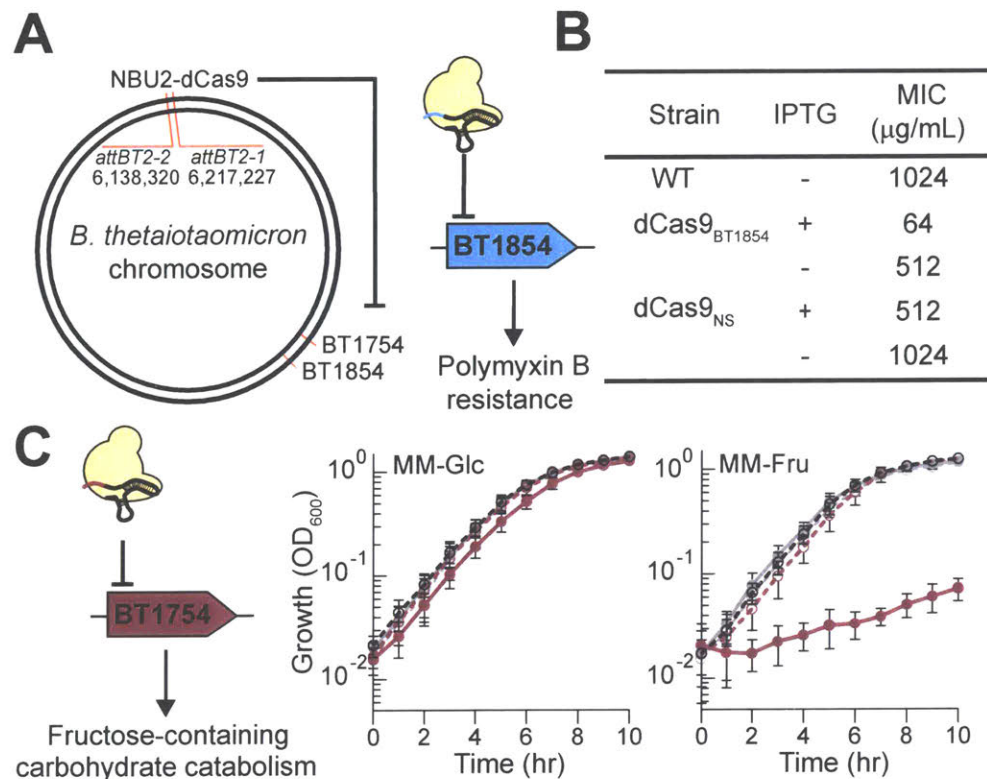
Figure 3-11: **CRISPRi-mediated repression of endogenous genes.** (A) Genomic location of endogenous genes targeted using CRISPRi. (B) Minimum inhibitory concentrations (MICs) of polymyxin B for cells with CRISPRi targeted against BT1854 (dCas9$_{BT1854}$) compared with wild-type (WT) cells or non-specific control cells (dCas9$_{NS}$). Reported values are the mode of three independent biological replicates made on three separate days. (C) CRISPRi was targeted against BT1754 (dCas9$_{BT1754}$). Growth curves of wild-type (WT) (black), dCas9$_{BT1754}$ (pink) or dCas9$_{NS}$ (gray) cells in minimal media supplemented with 0.5% glucose (MM-Glc) or 0.5% fructose (MM-Fru) in the presence (full line) or absence (dotted line) of 100μM IPTG. Error bars represent the standard deviation of three biological replicates made on different days.

3.7 *In vivo* Evaluation of *Bacteroides* Genetic Parts

We next investigated whether the function of our *B. thetaiotaomicron* genetic parts and modules can be maintained in the context of a complex microbiota. As wild-type strains of *Bacteroides* spp. are unable to stably colonize conventional specific-pathogen free (SPF) mice [235, 237], we employed an antibiotic regimen that promotes

B. thetaiotaomicron colonization without sterilizing the gut microbiota (Fig. 3-12A) [237, 238]. A ten-day treatment of animals with ciprofloxacin and metronidazole prior to bacterial inoculation was sufficient to maintain stable and high levels of colonization for the duration of the experiments (up to 12 days tested) (Fig. 3-12E-G).

Using this model, we tested the functionality of our inducible systems, CRISPRi, and integrases *in vivo*. First, SPF mice were colonized with the strain containing the arabinogalactan-inducible P_{BT0268} promoter driving expression of NanoLuc (Fig. 3-12E). Within a day of addition of arabinogalactan to the drinking water of the mice, luciferase activity in fecal pellets increased approximately 75-fold (Fig. 3-12B). Following removal of inducer from the drinking water, luciferase activity in the fecal pellets of mice fed inducer rapidly returned to baseline, demonstrating tight temporal control of gene expression dependent on arabinogalactan.

To investigate whether more complex genetic circuits perform in the context of the mouse microbiome, we evaluated the $dCas9_{NL3}$ repressor cascade, which is composed of the CRISPRi system as well as the P_{LacO23} IPTG-inducible promoter, within stably colonized *B. thetaiotaomicron*. Within 24 hours of adding IPTG to drinking water, CRISPRi elicited approximately a 20-fold reduction in gene expression compared to the uninduced control (Fig. 3-12C). The fold repression observed *in vivo* is similar to that measured *in vitro*. Luciferase activity returned to baseline 6 days following the removal of IPTG from drinking water. Moreover, expression of dCas9 and NanoLuc did not significantly impact *in vivo* fitness compared to uninduced controls (Fig. 3-12E and F).

To test the function of recombinases *in vivo*, we colonized mice with a *B. thetaio-taomicron* strain containing the rhamnose-inducible Int12 integrase memory switch (Fig. 3-9E). All mice were fed with plant-based chow that was determined to be composed of 0.3% rhamnose (w/w). In addition, after one day of colonization, the drinking water of half of the mice was supplemented with 0.5M rhamnose to further induce the memory switch. Stool was collected over the course of the experiment, and the absolute number of unflipped (wild-type) and flipped Int12 recognition sequences was determined by qPCR (Fig. 3-12D, Experimental Details). A background recom-

bination rate of ~11% per day was detected in mice fed on rhamnose-containing chow but not supplemented with rhamnose in their drinking water (Fig. 3-12D, "Chow"). In mice supplemented with exogenous rhamnose (Fig. 3-12D, "Chow + Rha"), the recombinase switch achieved >90% flipping in ≤1 day, a statistically significant increase over mice not supplemented with rhamnose in the water (p<0.01; Fig. 3-12D).
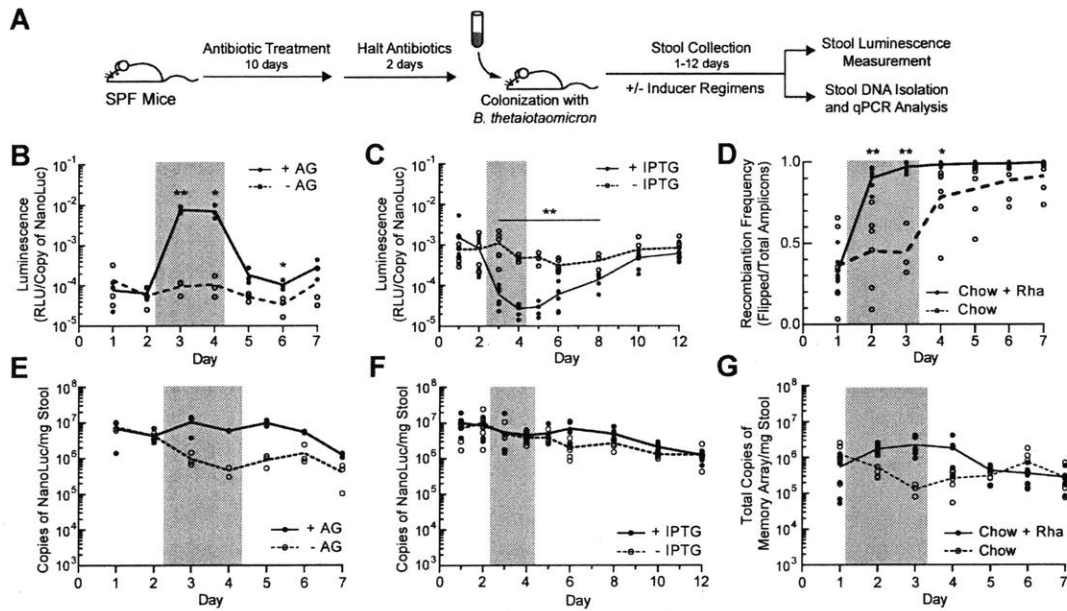
Figure 3-12: *In vivo* **function of genetic parts within** ***B. thetaiotaomicron*** **colonizing the mouse gut.** (A) Experimental timeline. Specific pathogen free (SPF) Swiss Webster mice were treated for 10 days with ciprofloxacin and metronidazole, and gavaged with *B. thetaiotaomicron* 2 days after cessation of treatment. Bacteria were administered on Day 0 (B-G). (B-C) Luciferase activity in fecal pellets of mice inoculated with strains possessing the arabinogalactan (AG) inducible P_{BT0268} (B) or IPTG-inducible CRISPRi $dCas9_{NL3}$ (C) systems. Mice were provided drinking water supplemented with 5% arabinogalactan (B: solid line), or 25mM IPTG (C: solid line) for 2 days after stool collection on Day 2 (grey box), or were maintained on normal drinking water throughout the entire experiment (dashed lines). Inducer water was removed on Day 4 after stool collection. Luminescence values were normalized to cell density as determined by qPCR using NanoLuc-specific primers. (D) SPF mice were colonized with *B. thetaiotaomicron* containing the rhamnose-inducible integrase construct P_{BT3763}-rpiL*C51-Int12. All mice were exposed to 0.3% rhamnose (w/w) in the plant-based chow. In addition, half of the mice had their drinking water supplemented with 500mM rhamnose after stool collection on Day 1 ("Chow + Rha", solid line) while the other half of the mice were maintained on normal drinking water throughout the entire experiment ("Chow", dashed line). Mice receiving rhamnose-supplemented water on Days 1 and 2 (grey box) were returned to normal water on Day 3 after stool collection. Recombination ratios were determined for fecal DNA as described in Figure 3F. For day 3 "Chow" samples, n=3. For all other days, n=6 for both treatment groups. (E-G) Cell densities of the arabinogalactan-inducible P_{BT0268}, (E) the $dCas9_{NL3}$ (F), or the rhamnose-inducible integrase (G) strains in the fecal pellets of inoculated mice. Grey boxes indicate the period of time over which mice were exposed to inducer in their drinking water. Bacterial loads were quantified by analyzing DNA extracted from fecal pellet using qPCR. The number of cells was determined using construct-specific primers and a standard curve generated with purified construct amplicons. Results were normalized to the weight of fecal material analyzed. For B-G, individual points represent independent biological replicates and the line represents the mean of the group. *P<0.05; **P<0.01.

3.8 Discussion

Here, we developed a versatile set of genetic technologies for the manipulation of the abundant gut symbiont *B. thetaiotaomicron*. Given the distinct transcriptional and translational control mechanisms in *B. thetaiotaomicron*, we laid the foundation for circuit design in *B. thetaiotaomicron* by identifying promoters and RBSs for this organism through literature searches and library construction. The resulting toolbox expanded upon the number and expression range of genetic parts available for *Bacteroidetes* (previous range: 10^2) [207, 208, 213, 214] and achieved ranges of expression

similar to those of libraries characterized for other gut-associated bacteria, including *E. coli* (range: 10^4-10^5) [239–241] and lactic acid bacteria (range: 10^3) [242, 243]. Using a similar approach, we characterized 4 orthogonal inducible expression systems. This library of *B. thetaiotaomicron* promoters and RBSs enabled us to import host-independent machinery, dCas9 and serine integrases, to implement more complex circuits. We envision that similar approaches, coupled with continuing advances in DNA sequencing and synthesis technologies, will enable the rapid and large-scale porting of genetic parts and devices from model organisms into new chassis that can address real-world applications. Lastly, we demonstrated the functionality of these genetic parts in the context of the mouse microbiome. Future work will focus on scaling the complexity of genetic circuits that can be implemented in commensal bacteria and pursuing long-term and quantitative control of microbial and mammalian host biology *in vivo*. For microbiome engineering applications, the ability to precisely modulate gene expression in commensal organisms should enable functional studies of the microbiome, non-invasive monitoring of *in vivo* environments, and long-term targeted therapeutics. For example, our constitutive and inducible systems, integrases, and CRISPRi regulators could be integrated for higher-order computation in *B. thetaiotaomicron*, such as digital logic, memory, and analog circuits. These engineered commensals could be used to map the dose-dependent and temporal effects of specific surface polysaccharides [244] or heterologous pathways [245] on colonization and maintenance of the gut microbiota and on host health. Higher-order combinations of inducible promoters linked with integrases could achieve Boolean logic with embedded cellular memory, enabling surveillance of the gut environment. Furthermore, environmental sensing with analog circuits coupled with precision expression control of heterologous pathways in *B. thetaiotaomicron* could be exploited for on-demand, localized, and rheostatic delivery of therapeutic molecules. We have also shown that the CRISPRi system can be used to dynamically manipulate bacterial processes in *B. thetaiotaomicron* by targeting endogenous genes. dCas9-mediated repression could be induced in a commensal library of *B. thetaiotaomicron* harboring distinct guide RNAs to identify genes required for *B. thetaiotaomicron* maintenance

[236] or interspecies interactions [246]. With these genetic resources, we envision that *B. thetaiotaomicron* will be a useful platform for cellular sensing, computation and actuation at the host-microbe interface in the gut.

3.9   Experimental Details

Bacterial strains and culture conditions.

B. thetaiotaomicron VPI-5482 (ATCC 29148) (Accession: AE015928.1) was used for all studies except for the integrase-based memory units. *B. thetaiotaomicron* VPI-5482 Δtdk was used for the insertion of the memory array and the characterization of integrases. *E. coli* was routinely grown in Luria Bertoni (LB) broth (Difco) supplemented with carbenicillin ($100\mu g$/mL) for plasmid selection. *B. thetaiotaomicron* was grown in brain heart infusion media (BHIS) (Difco), trypticase-yeast extract-glucose (TYG) media or minimal media (MM) supplemented with 10mg/L hemin (Sigma), 0.5g/L cysteine and 1mg/L Vitamin K3 [247]. TYG media contained: 10g/L trypticase, 5g/L yeast extract, 1g/L Na_2CO_3, 10mM glucose, 80mM potassium phosphate buffer (pH 7.3), 20mg/L $MgSO_4 \cdot 7H_2O$, 400mg/L $NaHCO_3$, 80mg/L NaCl, 0.0008% $CaCl_2$, and $4\mu g$/mL $FeSO_4 \cdot 7H_20$. MM media contained: 1g/L NH_4SO_4, 1g/L Na_2CO_3, 80mM potassium phosphate buffer (pH7.3), 900mg/L NaCl, 26.5mg/L $CaCl_2 \cdot 2H_2O$, 20mg/L $MgCl_2 \cdot 6H_2O$, 10mg/L $MnCl_2 \cdot 4H_2O$, 10mg/L $CoCl_2 \cdot 6H_2O$, $5\mu g$/L Vitamin B12, $40\mu g$/mL $FeSO_4 \cdot 7H_2O$, and 0.5% of carbon source. Cultures were grown and manipulated in a Coy anaerobic chamber with an atmosphere of 85% N_2, 5% H_2 and 10% CO_2 at 37 °C. All media was pre-reduced overnight in anaerobic atmosphere before inoculation of cultures. Antibiotics were added as appropriate: erythromycin (Em, $25\mu g$/mL), gentamicin (Gm, $200\mu g$/mL) and 5-fluoro-2'-deoxyuridine (FUdR, $200\mu g$/mL).

Genetic parts and plasmids.

Genetic parts and plasmids used in this study are listed in Table S3 and S4. Single guide RNAs (sgRNA) used in this study are listed in Table S5. pNBU2 constructs were conjugated into *B. thetaiotaomicron* from *E. coli*. The DNA memory array cassette [219] was introduced into the chromosome of a *B. thetaiotaomicron* VPI-5482 Δ*tdk* strain by allelic exchange.

pNBU2 constructs were conjugated into *B. thetaiotaomicron* with *E. coli* S17 λ*pir*, which contains the conjugative machinery of the plasmid RP4 integrated into its chromosome [248] or through tri-parental mating with *E. coli* DH10B λ*pir* and *E. coli* HB101 / pRK600 [249], which encodes the RP4 conjugative machinery. The pNBU2 plasmid encodes the *intN2* tyrosine integrase, which mediates sequence-specific recombination between the *attN* site of pNBU2 and one of two *attBT* sites located in the 3' ends of the two tRNA^{Ser} genes, BT_t70 and BT_t71, on the *B. thetaiotaomicron* chromosome (Fig. 3-2). Insertion of the pNBU2 plasmid inactivates the tRNA^{Ser} gene, and simultaneous insertion into both BT_t70 and BT_t71 is unlikely due to the essentiality of ^{Ser}. pNBU2-based vectors have been used for single-copy complementation in *B. thetaiotaomicron* in *in vitro* studies [250] and *in vivo* mouse models [199]. For S17 λ*pir* mating, an overnight culture of S17 λ*pir* was washed once with PBS and added to *B. thetaiotaomicron* at a ratio 1:5 (v/v). The mating mixture was pelleted, resuspended in 0.2mL of BHIS, spotted onto a BHIS agar plate and incubated upright in a 37°C incubator aerobically. For tri-parental mating, overnight cultures of DH10B λ*pir*, HB101 / pRK600, and *B. thetaiotaomicron* were subcultured, grown to an OD$_{600}$ of ~0.15-0.25, washed once with BHIS, and mixed in a 1:1:10 ratio. Tri-parental mating mixtures were spot plated as described above. After overnight incubation, cells from S17 λ*pir* or tri-parental matings were collected by scraping, resuspended in 3mL of BHIS, plated as serial dilutions on BHIS+Gm+Em, and incubated anaerobically for 48 hours at 37°C. Resultant colonies were re-isolated on BHIS+Gm+Em plates before further experimentation.

The DNA memory array cassette [219] was introduced into the chromosome of a

B. thetaiotaomicron δ*tdk* strain by allelic exchange as previously described [251]. The memory cassette was inserted between convergently transcribed genes BT2113 and BT2114 (between nt 2,660,694 and 2,660,695) to decrease the likelihood of disrupting native promoters or multicistronic messages. DNA fragments 3.5 kb in length were amplified from 5' and 3' of the insertion site and were cloned, flanking the memory array, into the counterselectable suicide plasmid pExchange-*tdk*, which encodes thymidine kinase. The resulting plasmid (pAT840) was conjugated into *B. thetaiotaomicron* δ*tdk* and cells were plated on BHIS+Em to select for strains in which pAT840 had recombined into the *B. thetaiotaomicron* δ*tdk* chromosome. The *tdk* allele encoded in the pExchange-*tdk* plasmid results in a 5'fluoro-2-deoxyuridine sensitivity (FUdR^S). Erm^R transconjugants were streaked on BHIS+FUdR to promote loss of the integrated plasmid. Isolated colonies were screened on BHIS for Erm^S+ FUdR^R phenotypes, and the presence of the memory array was confirmed by PCR and sequencing.

Construction of promoter and RBS libraries.

Promoters and RBS libraries were constructed by PCR using primers with specific or degenerate sequences.

Luciferase assay.

For the inducible promoter and dCas9 assays, *B. thetaiotaomicron* were lysed by adding 20 μl PopCulture reagent (EMD Millipore) supplemented with lysozyme solution (0.113% v/v) to 200 μL of cultures. For analysis of constitutive promoter and RBSs in *B. thetaiotaomicron*, 500 μL *B. thetaiotaomicron* cells were harvested by centrifugation at 4500 x g for 10 minutes and lysed by resuspending the cell pellet in 50 μL 1X BugBuster (Novagen) in 1X phosphate-buffered saline. NanoLuc luciferase assay (Promega) was performed according to the manufacturer's suggestions using 10 μL of *B. thetaiotaomicron* cell lysate and 10 μL NanoLuc Reaction Buffer. NanoLuc luciferase [252] is a small (19 kDa) modified shrimp luciferase. Efforts to use

members of the green fluorescent protein family [253] and a FMN-based fluorescent reporter [254] were not successful. NanoLuc oxidizes the exogenously-added substrate furimazine to produce glow-type bioluminescence ($E_{max} = 460$nm) with a signal half-life of 2 hr. By comparison, bacterial luciferase LuxAB (79 kDa) exhibited rapid signal decay when used to characterize gene expression in *Bacteroides* [255]. Luciferase has been used to characterize parts in *Bacillus subtilis* [256] and as a reporter in Gram-positive and Gram-negative bacteria [257].

Luciferase activities were measured with an integration time of 1 second at a gain setting of 100 in a BioTek Synergy H1 Hybrid Reader. Luciferase values were normalized to colony forming units (CFU) by measuring the OD_{600} of 200 μL of culture at the time of harvest in a BioTek Synergy H1 Hybrid Reader. The OD_{600} values were converted to CFU/mL using standard curves (OD_{600} vs. CFU/ml). Standard curves were generated by serial dilution and OD_{600} measurements ($OD_{600} = 0.015$ - 1.2) over 8 hours of growth.

Luciferase acivity in fecal pellets was monitored daily. Pellets were homogenized in 1 mL of PBS with a single 5mm stainless steel bead using a TissueLyser II (Qiagen). Samples were centrifuged at 500 x g for 30 seconds to pellet the bead and fibrous material. The luciferase assay was performed as described above using 25 μL of fecal homogenate and 25 μL NanoLuc Reaction Buffer.

Construction and analysis of the rpiL* RBS library.

The rpiL* RBS was mutagenized by amplifying the plasmid pAT593 (encoding P_{BT1311}-rpiL*-NanoLuc in a pNBU2 vector) with primer oAT617 in combination with either primer oAT614, primer oAT615, or primer oAT616 (Table S6), which were each synthesized with degenerate nucleotides at three positions ($4^3 = 64$ potential members in each library). The resulting fragments were digested with BbsI, ligated with T4 ligase, and digested with DpnI, yielding three scarless P_{BT1311}-rpiL* RBS libraries.

Growth and induction.

For inducible promoter and dCas9 assays, overnight cultures were diluted 1:100 in fresh TYG supplemented with inducer. The inducers used were IPTG (Goldman), L-rhamnose (Sigma), chondroitin sulfate A from bovine trachea (Sigma), and (+)-arabinogalactan from larch wood (TCI). Cultures were grown anaerobically at 37°C to an optical density (OD_{600}) of ~0.4-0.8 (~5 hours) before assaying luciferase activity.

In vitro genetic memory assays.

To identify integrases that function in *B. thetaiotaomicron*, pNBU2-based plasmids containing P_{AM4}-rpiL*RBS-integrase combinations were conjugated into *B. thetaiotaomicron* Δtdk harboring a chromosomal copy of the memory array. Cells were plated on BHIS+Gm+Em to select for strains that contained the promoter-RBS-integrase plasmids. DNA was isolated from transconjugants were grown were assayed by PCR for inversion of spacer sequences in the memory array cassette as described [219]. Control primers specific for unflipped (wild-type) memory array were used to detect the wild-type orientation of the spacer sequences (Table S6). To characterize the function of integrase Int12 in *B. thetaiotaomicron*, a pNBU2-based plasmid containing P_{rha}-rpiL*C51RBS-Int12 was constructed and conjugated into *B. thetaiotaomicron* memory array strain described above. For the recombinase response curve, an overnight culture grown in TYG was diluted (1:100) into TYG or TYG supplemented with rhamnose and grown for 8 hours at 37°C. The final OD_{600} of each culture was measured using a Cary 50 UV-Vis spectrophotometer (Varion), and 500 μl of each cell culture was harvested by centrifugation, flash frozen, and stored at -80°C until DNA was extracted. For the recombinase time course assay, cells were grown in TYG supplemented with 10 mM rhamnose for 8 hours at 37°C, harvested periodically, and stored as described above. DNA was isolated using the DNeasy Blood and Tissue Kit (Qiagen). Levels of integrase-mediated recombination in vitro were quantified using qPCR with specific to the wildtype (unflipped) (oAT836/837) or recombined (flipped) (oAT836/838) DNA at the target sequence. For absolute quantitation, we

used a standard curve generated with purified DNA standards. Flipping ratio was calculated as (copies of flipped DNA) / (copies of flipped DNA + copies of unflipped DNA).

Polymyxin B susceptibility.

Minimum inhibitory concentrations (MICs) were determined by broth microdilution using TYG broth according to modified CLSI guidelines [258]. Overnight cultures were diluted 1:100 in TYG with or without 100 μM IPTG and grown until an OD_{600} of 0.5-0.7. Two-fold serial dilutions of polymyxin B starting at a concentration of 1024 μg/mL in TYG with or without 100μM of IPTG were prepared in a 96-well microtiter plate. Samples were added to the microtiter plate at a final concentration of 10^6 CFU/mL and incubated anaerobically at 37°C for 24 hours. The MIC was determined as the lowest concentration of polymyxin B that completely inhibited growth as measured by eye.

Fructose/glucose growth assay.

Overnight cultures were grown in TYG supplemented with or without 100μM of IPTG. The following morning, samples were diluted 1:100 in MM supplemented with either 0.5% glucose or 0.5% fructose as the sole carbon source and with or without 100μM IPTG. Cultures were incubated anaerobically at 37°C and 200 μL samples were withdrawn from the cultures every hour for 10 hours to monitor growth (OD_{600}) in a BioTek Synergy H1 Hybrid Reader.

Animals.

All animal study protocols were approved by the MIT Animal Care and Use Committee. Specific pathogen free (SPF) female Swiss Webster mice (5-8 weeks) were purchased from Charles River and were housed and handled in non-sterile conditions. Mice were fed irradiated mouse chow (PicoLab Rodent Diet 20, LabDiet) and provided autoclaved or sterile filtered water for the duration of the experiment. Prior to bacte-

rial gavage, mice were administered ciprofloxacin-HCl (0.625 g/L), metronidazole (1 g/L) in sugar-sweetened drinking water for 7 days and were subsequently treated with metronidazole (100 mg/kg) by oral gavage every 24 hours and ciprofloxacin (0.625 g/L) dissolved in drinking water for three days. Animals were transferred to a clean cage and provided fresh medicated water on the fifth day of the antibiotic regimen. Two days after the cessation of antibiotic treatment, animals were inoculated with engineered strains of *B. thetaiotaomicron* ($\sim$5x10^8 CFU/mouse) by oral gavage. Mice belonging to the same treatment group were cohoused for the duration of the experiment. Fecal pellets were collected daily for luciferase and qPCR analysis (described below).

Enumeration of *B. thetaiotaomicron* strains *in vivo*.

Strains were enumerated by qPCR with strain-specific primers (Table S6) targeting the NanoLuc (mmD662/663) or flipped (oAT836/838)/unflipped (oAT836/837) memory array alleles in total fecal DNA. For absolute quantitation, we used a standard curve generated with purified DNA standards.

Function of inducible promoters and CRISPRi *in vivo*.

After collection of stool 2 days post-bacterial gavage, mice were provided drinking water supplemented with 25mM IPTG, 5% arabinogalactan or kept on sterile water as controls. Normal drinking water was returned to animals 4 days post-bacterial gavage. For the arabinogalactan experiment, animals were transferred to a clean cage following removal of supplemented drinking water. For the CRISPRi experiment, animals were transferred to a clean cage 7 days post-bacterial gavage. Luciferase values were measured daily and normalized to the cell density by determining the copy number of NanoLuc by qPCR using primers mmD662/663.

Quantification of memory unit recombination *in vivo*.

After collection of stool on Day 1, mice were provided with drinking water supplemented with 0.5 M rhamnose or were maintained on sterile water as controls. Mice were transferred to clean cages and returned to normal drinking water 3 days post-bacterial gavage. Levels of integrase-mediated recombination *in vivo* were determined by analyzing fecal DNA (1 μl) by qPCR with primers as described for the *in vitro* analysis. Flipping ratio was calculated as (copies of flipped DNA) / (copies of flipped DNA + copies of unflipped DNA).

Determination of L-rhamnose content in the mouse chow

Rhamnose biosynthetic pathways are absent in higher vertebrates, but rhamnose is a common component of the plant and bacterial cell wall [259]. The concentration of rhamnose in the mouse chow was determined using a commercially available L-Rhamnose Assay Kit (Megazyme), according to manufacturer's instructions. Samples were prepared by heating 100mg of chow in 5 mL of 1.3M HCl at 100 °C for 1 hour. Hydrolysates were neutralized with NaOH and insoluble debris was pelleted by centrifuging for 10 minutes at 1500 x g.

Data analysis, statistics, and computational methods.

All data were analyzed using Graphpad Prism version 6.05 (GraphPad Software, San Diego, CA, USA, http://www.graphpad.com/). Error bars represent the standard deviation of 3 experiments carried out on different days. Fold repression was calculated by dividing the luminescence of uninduced samples (RLU/CFU) by the luminescence of samples treated with maximum inducer concentration. All response curves were fit to a Hill function: $Y = (B_{max}X_n)/(K_n+X_n)+C$, where X is the input, Y is the output, B_{max} is the maximum luminescence, n is the Hill coefficient, K is the threshold and C is baseline luminescence. RBS frequency logos were generated for the region of the rpiL* RBS targeted by mutagenesis by comparing the frequency of each nucleotide at each position within the subpopulation (strongest 7% of RBSs

and weakest 7% of RBSs) to the frequency of that nucleotide at that position in the full library.

Chapter 4

Engineering *Bacteroides* Outer Membrane Vesicles

Abstract

The microbiota forms an intimate link with the host immune system during development and plays an integral role in the maintenance of immune homeostasis. Outer membrane vesicles (OMVs) are membranous structures secreted by commensal organisms that can manipulate host immune responses. Anti-inflammatory in nature, OMVs derived from *Bacteroides* spp. could serve as potent agents to dampen immune responses. Using proteomic and bioinformatics approaches, we analyzed *B. thetaiotaomicron* OMV proteins and found SpII-type lipoproteins were enriched in the vesicles. Translational fusions to these native OMV proteins could tether heterologous proteins to vesicles and lipoprotein signal sequences were sufficient to mediate OMV incorporation in multiple *Bacteroides* spp. Next, we employed these *Bacteroides* Vesicle Incorporation Tags (BVITs) to package a model antigen to evaluate how *Bacteroides* OMVs impact antigen presentation and T cell activation. Antigen-loaded OMVs could be correctly processed and presented in a dendritic cell-T cell co-culture model and resultant T cells secreted high levels of anti-inflammatory cytokine. Based on these findings, we demonstrate a strategy to package heterologous cargo into *Bacteroides* OMVs to manipulate adaptive immune responses.

4.1 Introduction

A constant dialogue between the tens of trillions of microbes [260] and their human hosts is the bedrock of a healthy gut immune system. Molecular components of commensal organisms serve an integral role in the development of immune responses. Early life exposure to poorly immunogenic lipopolysaccharide can lead to inhibited endotoxin tolerance and autoimmune disease [56]. Additionally, protein antigens derived from segmented filamentous bacteria can induce Th17 cells, but, when heterologously expressed in *Listeria monocytogenes*, induce Th1 responses, suggesting that the context in which bacterial molecules are presented to the immune system can significantly impact outcomes [64]. Certain bacterial taxa that synthesize zwitterionic capsular polysaccharides can induce IL-10 producing regulatory T cells that dampen immune responses [58, 261]. The diverse set of immune responses that can be elicited by microbially-derived molecules positions the microbiota as central node to manipulate the host immune system.

Outer membrane vesicles (OMVs) serve as a major line of communication between microbe and host. OMVs are 20-250nm, spherical proteoliposomes secreted by Gram negative bacteria, consisting of outer membrane and periplasmic content [262] [263] (Fig. 4-1). Phospholipids, lipopolysaccharide, integral and peripheral membrane proteins, peptidoglycan and periplasmic proteins are packaged into OMVs and secreted to distal sites. The mechanism of OMV biogenesis has yet to be fully elucidated and may be either stochastic or regulated [264]. Vesicle budding tends to occur in destabilized membrane regions with reduced crosslinking to peptidoglycan or increased lipopolysaccharide charge repulsion [264]. Increased vesiculation is observed as a response to envelope stress, with misfolded proteins packaged as cargo in the membrane blebs [265]. Immunogenic proteins, tissue proteases, antibiotic resistance enzymes, toxins, and other virulence factors are enriched in OMVs derived from pathogenic bacteria, suggesting they play an important role in microbe-host interaction [266] [267]. In *B. fragilis*, anti-inflammatory polysaccharide A is ferried to the immune system on OMVs [268, 269]. In *B. thetaiotaomicron*, sulfatase-containing

OMVs can digest mucosal polysaccharides, transit through the epithelial layer, and directly interact with immune cells residing in the lamina propria [270]. Based on this intimate OMV-mediated immune system interaction, engineered OMVs could serve as a potent, targeted immunomodulatory agents. Indeed, *Escherichia coli* OMVs have been engineered to package heterologous protein and polysaccharide cargo for the development of vaccines [271, 272]. The immunogenic properties of *E. coli* OMVs act as an adjuvant to amplify immune response to packaged antigens. However, certain *Bacteroides* spp. elicit anti-inflammatory reactions in the host through the induction of regulatory T cell responses [58, 59, 273] and could thus potentially mediate immunological tolerance to packaged cargo.
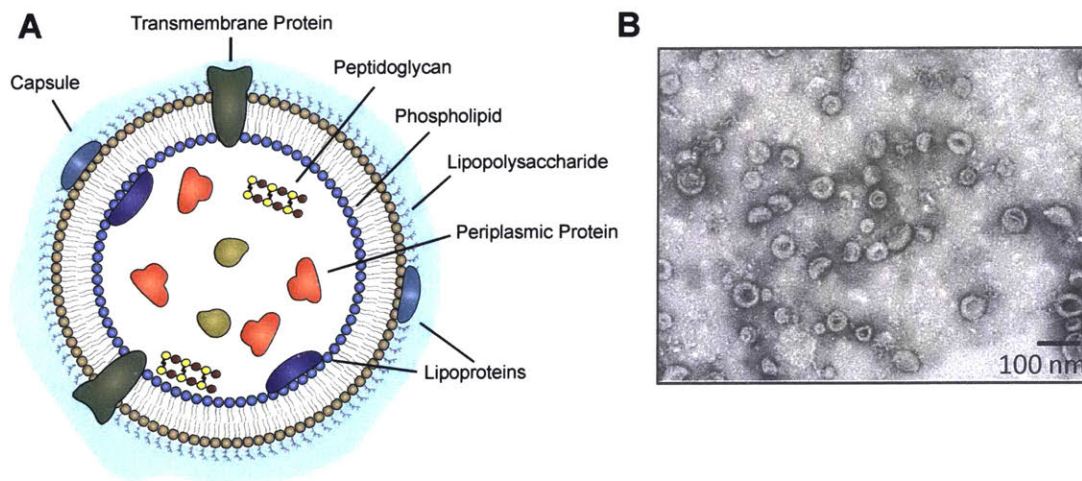
Figure 4-1: **Outer membrane vesicles:** (A) Schematic of outer membrane vesicles. (B) Transmission electron micrograph of outer membrane vesicles isolated from *B. thetaiotaomicron.*

Here, we describe a strategy for heterologous protein incorporation into OMVs derived from *Bacteroides* spp. Proteomic analysis of *Bacteroides* OMVs revealed an enrichment of SpII-type lipoproteins, suggesting that conserved secretion sequences could serve as *Bacteroides* Vesicle Incorporation Tags (BVITs). Sequential truncations and bioinformatics analysis led to the identification of a set of BVITs that direct protein packaging into OMVs in *B. thetaiotaomicron, B. fragilis* and *B. vulgatus.* Using these sequence tags, we explore fusion of mammalian cytokines to OMVs

as a means to modulate the host immune system. Subsequently, we demonstrate that OMVs containing the model antigen ovalbumin (Ova) can be processed and presented in dendritic cell-T cell co-culture models. Engineered *Bacteroides* OMVs could potentially serve as an interesting means to both deliver proteins to the host immune system and to study commensal antigen presentation in the gut mucosa.

4.2   Identification of OMV Proteins

Although the sorting mechanism is unknown, proteomic analyses of OMVs suggest that specific outer membrane and periplasmic proteins are packaged in a non-random manner [274]. In order to identify the molecular determinants that govern protein incorporation into *Bacteroides* OMVs, we performed a proteomic survey of vesicles purified from *B. thetaiotaomicron* cultures. The analysis of tryptically digested OMVs yielded 37 unique proteins present in the vesicle fraction, 26 of which had previously been identified in mass spectrometry analyses of *Bacteroides* OMV contents [275] (Appendix). Homology-based annotation of identified proteins revealed that OMV proteins are largely uncharacterized or involved in macromolecule catabolism. 'SusD-like' proteins, belonging to the 'starch-utilization system' family, were particularly enriched in the OMVs, supporting the notion that OMVs serve as a public good that could aid in population-wide metabolism of complex substrates.

The majority (73%) of identified proteins contain signal peptidase II (SpII)-type cleavage sites in their N-terminus, suggesting that lipoproteins are enriched in OMVs. Lipoproteins are post-translationally modified proteins that are peripherally anchored into the bacterial cell envelope [276, 277]. Lipoprotein signal sequences characteristically possess a positively charged N-terminus, a central hydrophobic region and a polar C-terminal region containing an SpII cleavage site or 'lipobox.' Preprolipoproteins are exported into the periplasm through traditional secretion mechanisms where it is process by SpII. Upon cleavage into the mature peptide, a conserved N-terminal cysteine residue is acylated on its sulfhydryl group, such that the protein is anchored into the membrane. For proteins destined for the outer membrane, the LolABCDE

complex mediates transport of mature lipoproteins across the aqueous periplasm and into the inner or outer leaflet of the outer membrane.
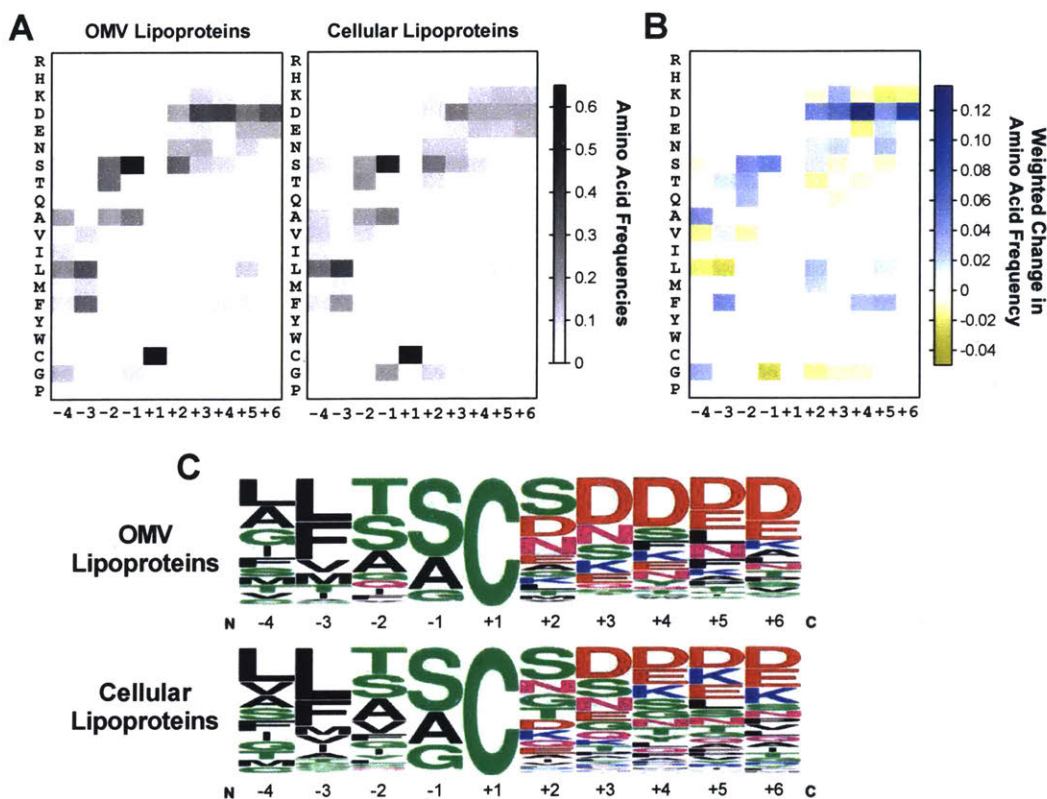
Figure 4-2: **Sequence Profile of *Bacteroides* Lipoproteins:** (A) Amino acid frequency of *Bacteroides* lipoproteins relative to the conserved cysteine residue. OMV lipoproteins represent those identified in proteomic screens (Appendix and [275]). Cellular lipoproteins consist of all predicted *B. thetaiotaomicron* and *B. fragilis* lipoproteins excluding OMV lipoproteins. (B) Differences in amino acid frequency at various positions in OMV lipoproteins as compared to cellular lipoproteins. The score was calculated as the log$_2$-fold change of amino acid frequency weighted by the frequency of a given amino acid in OMV lipoproteins. (C) Sequence logos of OMV and cellular lipoproteins.

In order to determine if OMV-associated lipoproteins possess unique SpII sequence signatures compared to cellular lipoprotein counterparts, we analyzed frequencies of amino acids in predicted lipoprotein signal sequences (Fig. 4-2A). Using LipoP [278] and SignalP [279], we compiled all polypeptides predicted to encode SpII lipoproteins from the reference genomes of *B. thetaiotaomicron* and *B. fragilis*. Lipoproteins that

were not identified in proteomic analyses of OMVs were considered to be cellular lipoproteins (n=969 for cellular lipoproteins; n=100 for OMV lipoproteins). While the residues N-terminal to the modified cysteine form a conserved lipobox, the residues following the cysteine are much more degenerate in composition (Fig. 4-2B). As compared to the conserved Gram positive and negative lipobox (L_{-3}/A(S) $_{-2}$/G(A) $_{-1}$/C_{+1}) (Zückert, 2014), *Bacteroides* possess the more variable L(FMVI) $_{-3}$/TSA(V) $_{-2}$/S(AG) $_{-1}$/C_{+1} consensus sequence (Fig. 4-2C). OMV lipoproteins tended to favour serine as opposed to glycine at -1; serine, threonine and glutamine as opposed to valine at -2; and phenylalanine to leucine at -3. Moreover, OMV lipoproteins tended to contain more aspartic acid and asparagine residues in the +2 to +6 positions as compared cellular lipoproteins which were more glycine and lysine rich. Interestingly, in Proteobacteria, sorting of lipoproteins to the outer or inner membrane is dictated by the identity of amino acids in the +2 or +3 positions [280, 281]. In *E. coli*, aspartic acid in the +2 position leads to retention in the cytoplasmic membrane. The differential abundance of aspartic acid in +2 in *B. thetaiotaomicron* and *B. fragilis* OMV lipoproteins suggests that lipoprotein sorting signals differ from those in Proteobacteria.

Based on the relatively minor differences in individual amino acid frequencies in OMV and cellular lipoproteins, OMV incorporation does not seem to depend on simple primary amino acid sequences. Subtle motifs that are not as immediately apparent in a global bioinformatics analysis or secondary or tertiary structure differences more than likely govern enrichments of lipoproteins in OMVs.
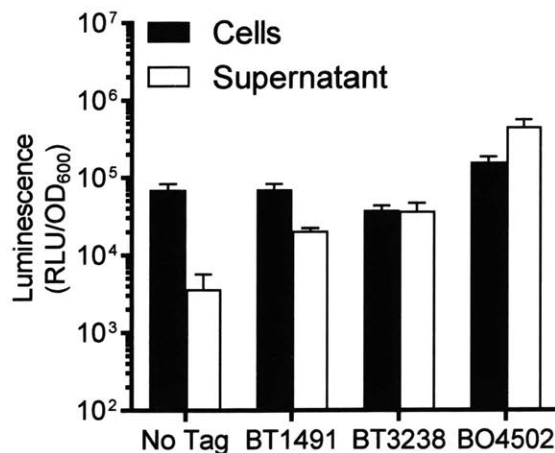
Figure 4-3: **Fusion of NanoLuc to OMV Proteins Leads to Enrichment in Culture Supernatant:** Normalized luminescence activity of cell pellets and supernatants of *B. thetaiotaomicron* strains expressing un-tagged NanoLuc (No Tag) or N-terminal fusions of full length BT1491, BT3238, and BO4502. Error bars represent standard deviation of three independent biological replicates (n=3).

4.3 Development of *Bacteroides* Vesicle Incorporation Tags (BVITs)

4.3.1 OMV-Associated Lipoproteins as Anchors for Fusion Proteins

Proteomic and bioinformatics analyses yielded a collection of proteins enriched in *Bacteroides* OMVs. We reasoned that these natural OMV lipoproteins can serve as anchors for heterologous incorporation. We chose two initial proteins as candidate OMV anchors: BT1491, hypothetical protein, and BT3238, a SusD-homolog. Both of these proteins were identified in our and previous proteomic assays and were strongly expressed on Coomassie-strained gels of total OMV proteins [275]. As a control, we also included the *B. ovatus* inulinase BO4502, which was substantially enriched in *B. fragilis* OMVs when heterogously expressed, suggesting a universal sorting mechanism in *Bacteroides* spp. [275] C-terminal fusions of NanoLuc luciferase to these full-length OMV proteins led to enrichment of luciferase activity in *B. thetaiotaomicron*

supernatants compared to an untagged control, suggesting that packaged cargo is secreted along with the OMV proteins (Fig. 4-3). In contrast, luciferase activity in cell pellets did not substantially increase in NanoLuc fusions. To confirm fusions to OMV anchor led to vesicular packaging, we purified OMVs from *B. thetaiotaomicron*. Luciferase activity was elevated in fusion protein strains compared to intracellular controls (Fig. 4-4). Additionally, only fusion proteins could only be detected in purified OMVs by Western blot (Fig. 4-4). Fusion of heterologous cargo to native OMVs is therefore a viable strategy to rationally direct protein packaging into OMVs.
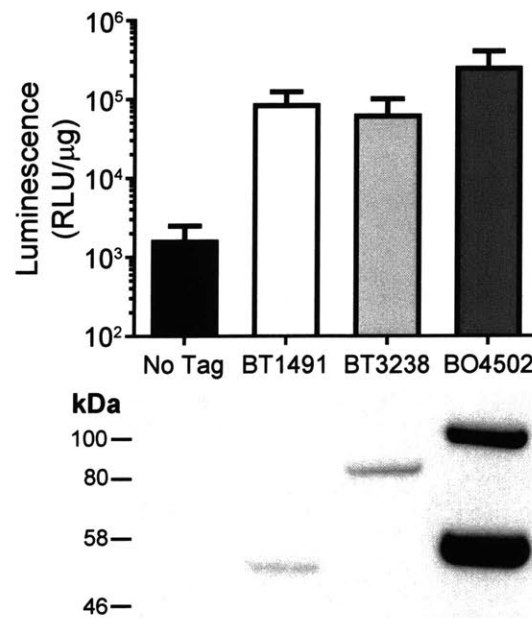
Figure 4-4: **OMV Protein-NanoLuc Fusions Are Enriched in Purified OMVs:** Luminescence activity of OMVs purified from *B. thetaiotaomicron* strains expressing un-tagged NanoLuc (No Tag) or N-terminal fusions of full length BT1491, BT3238, and BO4502. Error bars represent standard deviation of three independent biological replicates (n=3). Luminescence activity was normalized by the total protein content of OMVs. BT1491- (predicted MW: 47 kDa), BT3238- (predicted MW: 76 kDa), and BO4502- (predicted MW: 91kDa) fused NanoLuc was detected by Western blot. 4 μg of OMV total protein was analyzed. Note that BO4502-NanoLuc fusions yield a low molecular weight band that is likely a cleavage product of the mature protein. Blot is representative of three independent experiments.

4.3.2 Serial Truncations of Lipoproteins to Identify BVITs
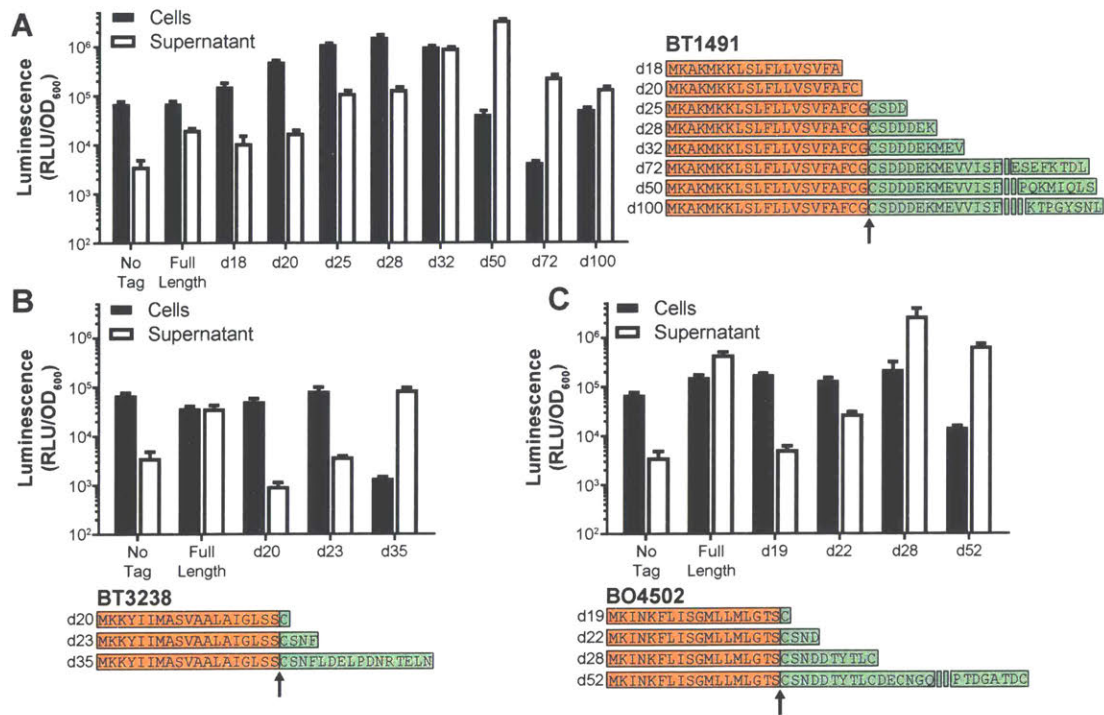
Figure 4-5: **Sequential Truncation of Lipoprotein Signal Sequences is Sufficient to Mediate Secretion of Heterologous Cargo:** The N-termini of BT1491 (A), BT3238 (B), and BO4502 (C) were serially truncated and fused with NanoLuc luciferase. OD$_{600}$-normalized luminescence activity in *B. thetaiotaomicron* cell pellets and culture supernatants was measured. Amino acid sequences of truncated variants are depicted with the cleaved pre-protein signal sequence in orange and the N-terminus of the mature protein in green. Arrow represents SpII cleavage site. Error bars represent standard deviation of three independent biological replicates (n=3).

Although full-length native proteins could serve as anchors for heterologous incorporation of cargo into OMVs, overexpression of functional protein could have deleterious effects on bacterial fitness, community dynamics, or host health. We therefore undertook an empirical approach to identify a minimal sequence that is sufficient to direct OMV cargo packaging. Serial N-terminal truncations of BT1491 revealed that a sequence as short as 25 amino acids was sufficient to lead to increased luminescence in culture supernatants (Fig. 4-5). Increasing the BVIT length to 32 amino acids increased secretion further. As compared to predicted lipoprotein sequences, the BVIT corresponded to an additional 9 amino acids past the lipobox, consistent

with observations that N-terminal residues in the mature protein are important for export [277]. A similar sequential truncation approach was applied to BT3238 and BO4502 to identify minimal BVITs, which were 35 and 28 amino acids, respectively (Fig. 4-5). Interestingly, as BVIT secretion efficiency increased, the intracellular luminescence decreased, suggesting that BVITs are an efficient means to shed protein from the cell.

BVITs were sufficient to direct incorporation into purified OMVs derived from *B. thetaiotaomicron* (Fig. 4-6A-B). OMVs contained equivalent or higher luminescence than full-length protein fusions. To determine the localization of cargo in OMVs, we performed a proteinase K protection assay. Proteins localized with the OMV lumen should be protected from proteinase K digestion, whereas those displayed on the surface would be susceptible. Surface exposure of certain cargo may be important for interaction with host cell-surface receptors or to metabolize low permeability substrates. BVITs derived from BT1491, BT3238 and BO4502 were all susceptible to proteinase K digestion, suggesting these lipoproteins are natively surfaced exposed (Fig. 4-6C). BVITs are thus sufficient to mediate packaging of heterologous proteins in OMVs, such that cargo is surface exposed.

4.3.3 Extension of BVIT Design Rules

Serial truncations of OMV proteins suggested that *in silico* predicted lipoprotein signal sequences could inform the development of BVITs. An additional 3-15 amino acids C-terminal to the predicted acylated cysteine residue were sufficient to mediate OMV packaging (Fig. 4-7). Using this design rule, we develop BVITs based on additional lipoproteins identified in proteomics screens: BF1504, BF1539, BF3342, and BT1955. C-terminal fusions of NanoLuc to BVITs derived from these lipoproteins led to enrichment of luminescence activity in culture supernatant (Fig. 4-7B) and purified OMV fractions (Fig. 4-7C) compared to un-tagged NanoLuc. Next, we investigated if these BVITs, originally derived from *B. thetaiotaomicron*, *B. fragilis* and *B. ovatus*, could mediate heterologous cargo incorporation into OMVs in other *Bacteroides* spp. When expressed in *B. fragilis*, and *B. vulgatus*, BVIT-tagged NanoLuc was enriched
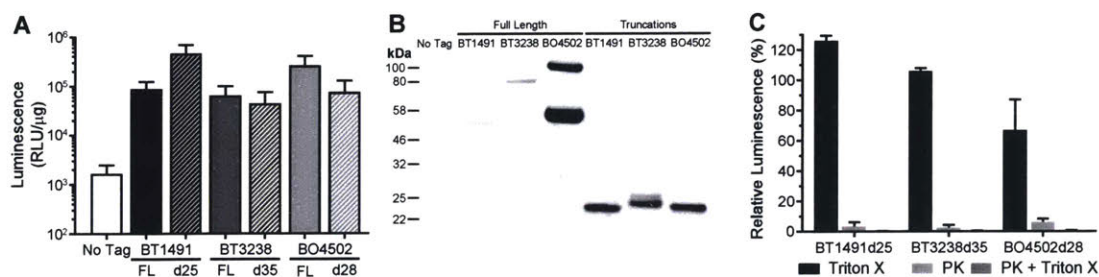
Figure 4-6: **Truncated Lipoproteins Are Sufficient for OMV Incorporation:** (A) Luminescence activity of full length (FL) or truncated lipoprotein OMV anchor fusions to NanoLuc was measured in OMVs purified from *B. thetaiotaomicron*. Luminescence values were normalized to total protein content of OMVs. (B) Western blot analysis of *B. thetaiotaomicron* purified OMVs containing full length (FL) or truncated lipoprotein fusions to NanoLuc. 4 μg of purified OMVs were separated by gel electrophoresis and were probed with anti-His. Note that the first 4 lanes of the blot are identical to those in Figure 4-4. Predicted MW: BT1491d25 – 21.1 kDa; BT3238d35 – 22.2 kDa; BO4502d28 – 21.5 kDa. (C) Purified OMVs isolated from *B. thetaiotaomicron* strains expressing BT1491d25-NL, BT3238d35-NL, and BO4502d28-NL were subject to proteinase K digestion for 9 hours at 37 °C. Luminescence activity was measured and compared to activity in non-protease treated controls. Triton-X was included as a positive control to disrupt OMVs to render proteins accessible to proteinase K digestion. Error bars represent standard deviation of three independent biological replicates (n=3).

in culture supernatants and purified OMVs. While most BVITs exhibit comparable efficacy in all species, BF1504 and BF3342 performed poorly in their native organism. Lipoprotein signal sequences are therefore an effective means to package cargo into OMVs, suggesting a common sorting mechanism across the *Bacteroides* genus.

4.3.4 Mutagenesis of BVITs to Modulate Secretion

Lipoproteins destined for OMV secretion must be efficiently trafficked to the outer membrane. The +2 and +3 amino acids of lipoprotein sequences have been implicated in localization within the cellular envelope [280, 281]. Therefore, we explored how primary amino acid sequence of BVITs impacts secretion efficiency through saturated mutagenesis of the +2 and +3 residues of BT1491d32 and BO4502d28 fused to NanoLuc (Fig. 4-8). As predicted, alterations of the amino acid sequence at these positions affected the proportion of intracellular or extracellular protein as compared
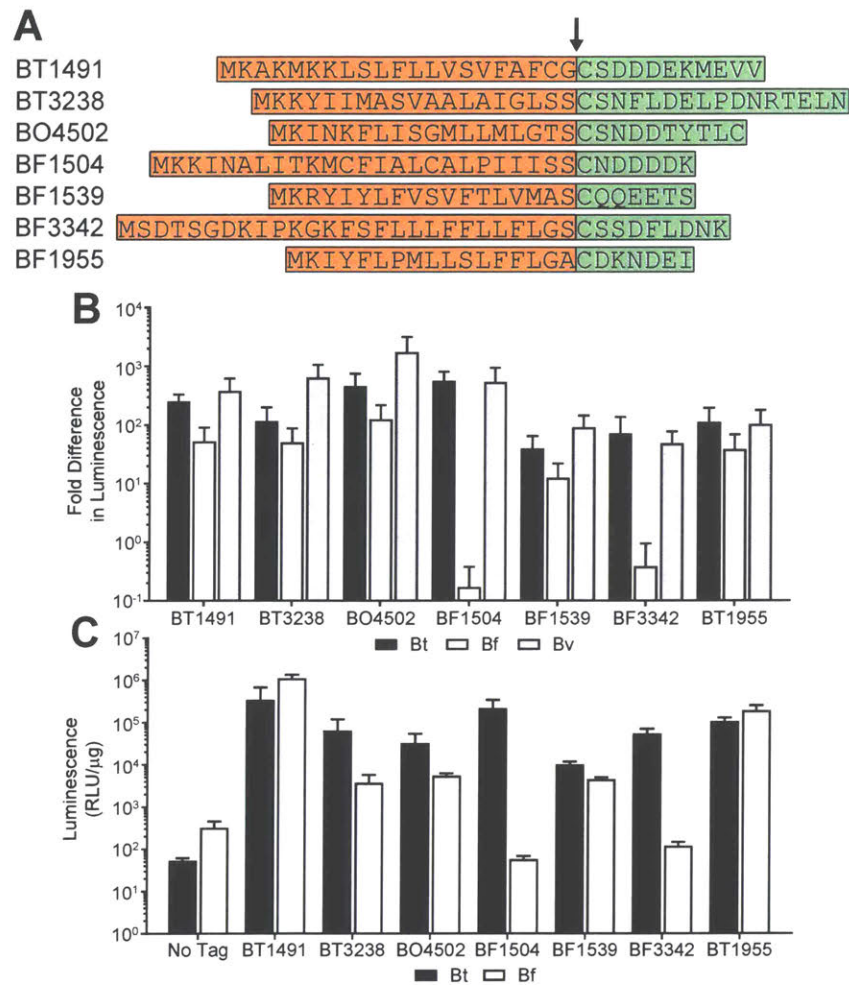
Figure 4-7: **BVIT Design Rules Are Applicable to Lipoproteins Across the _Bacteroides_ Genus:** (A) Primary amino acid sequences of BVITs derived from _B. thetaiotaomicron_ or _B. fragilis_ lipoproteins. Arrow indicates SpII cleavage site. (B) Fold change in luminescence in BVIT-tagged NanoLuc as compared to a non-tagged control in culture supernatants of _B. thetaiotaomicron_ (Bt), _B. fragilis_ (Bf), and _B. vulgatus_ (Bv). (C) Luminescence activity in OMVs purified from Bt or Bf expressing NanoLuc fused to various BVITs. Luminescence values were normalized by the OMV total protein content. Error bars represent standard deviation of three independent biological replicates (n=3).

to wild-type BVITs. In general, changes in luminescence activity in supernatants and pellets tended to be inversely related, such that increased secretion led to decreased cellular association. Basic amino acids, tyrosine and proline tended to decrease secretion and lead to higher cellular expression. Moreover, common residues identified in bioinformatics analyses (Fig. 4-3C) were consistent with mutagenesis results.
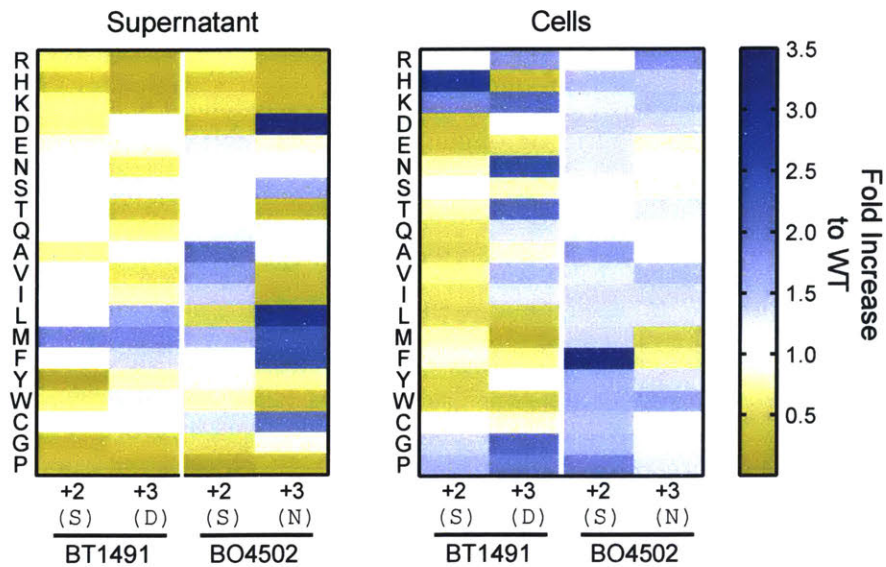
Figure 4-8: **Mutagenesis of BVITs to Modulate Secretion:** The +2 and +3 residues of BT1491d32 and BO4502d28 were substituted with all other amino acids and fused to NanoLuc reporter protein. OD_{600}-normalized luminescence activity of supernatants and cell pellets from *B. thetaiotaomicron* cultures expressing mutagenized or wild-type BVITs was measured. Data is expressed as the fold change in luminescence as compared to the wild-type BVIT sequences.

4.4 Incorporation of Immunomodulatory Cytokines into *Bacteroides* OMVs

With a strategy to package heterologous cargo protein into OMVs, we next sought to engineer *Bacteroides* OMVs to affect the host immune system. Interleukin-10 (IL-10) is a pleiotropic anti-inflammatory cytokine that is serves as a master negative regulator of immune responses [282]. IL-10 acts by binding to cell surface receptors and altering regulatory programs involved in inflammation in macrophages, T cells and B cells [282]. Importantly, deficiencies in IL-10 have been implicated in the incidence of inflammatory bowel diseases and loss-of-function mutations in IL-10 or its receptor is a risk factor [283]. Microbial expression of IL-10 has previously been explored as a treatment for colitis and *Lactococcus lactis* strains engineered to secrete IL-10 have been protective in dextran sodium sulfate murine models [94]. Likewise, IL-22 is a cytokine thought to be important in the resolution of gut lesions in inflam-

matory bowel disease [284, 285]. IL-22 acts on epithelial and goblets cells to stimulate proliferation to maintain intestinal barrier integrity and wound healing [284, 285]. In addition, ectopic expression of recombinant IL-22 in the gut has been demonstrated to reduce pathology in murine dextran sodium sulfate models of colitis [286]. Packaging mammalian cytokines onto *Bacteroides* OMVs could thus prove an effective means to engineer the gut immune system. The evidence that *Bacteroides* OMVs can transit into the lamina propria through the mucosa [270] suggests that OMVs could act as vessels to deliver cytokines to host immune cells. Moreover, as cargo is surface exposed (Fig. 4-6C), OMV-packaged cytokines would be accessible to binding their cognate receptors.
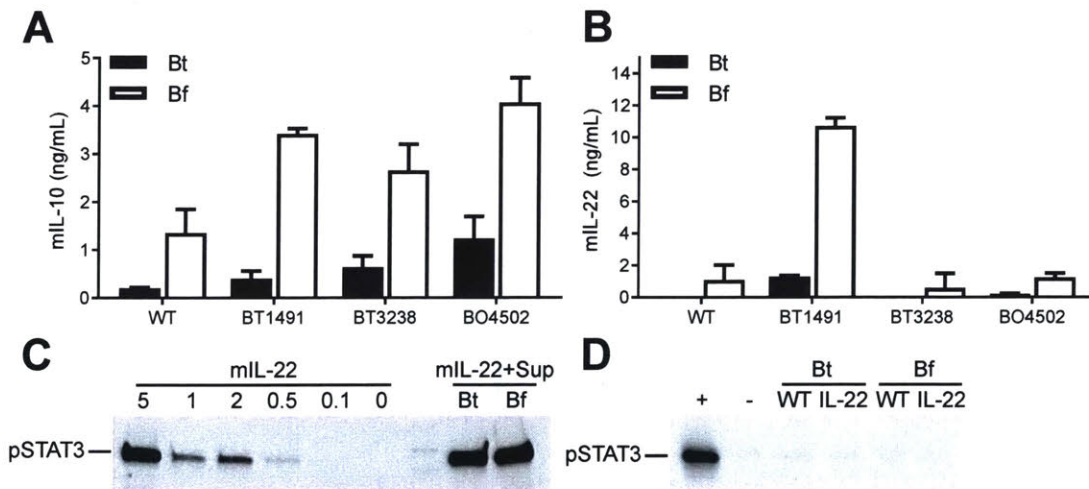
Figure 4-9: **Packaging Mammalian Cytokines into *Bacteroides* OMVs:** Levels of BVIT-fused mIL-10 (A) and mIL-22 (B) in *B. thetaiotaomicron* (Bt) and *B. fragilis* (Bf) culture supernatants was measured by ELISA. Error bars represent standard deviation of three independent biological replicates (n=3). (C) STAT3 phosphorylation in HT-29 cells exposed to various concentrations of recombinant mIL-22 was evaluated by Western blot. Recombinant mIL-22 was incubated in Bt or Bf supernatant to assay if a component of spent media inhibited IL-22 signaling through STAT3. (D) Bioactivity of BT1491-fused mIL-22 was assayed by verifying STAT3 phosphorylation in HT-29 cells. An equivalent of /sim5-10 ng/mL of mIL-22 was added in *Bacteroides* samples. 5 ng/mL of recombinant mIL-22 (+) was included as a positive control. Blots in C and D are representative of at least 3 separate experiments.

To incorporate murine IL-10 and IL-22 into *Bacteroides* OMVs, we fused these

cytokines to identified BVITs (Fig. A and B). Expression of BVIT-mIL-10 fusions led to the accumulation of ng/mL levels of cytokines in culture supernatants of both *B. thetaiotaomicron* and *B. fragilis* (Fig. 4-9A). Similarly, BVIT fusions yielded similar levels of mIL-22 in culture supernatants (Fig. 4-9B). Expression efficiency seemed to be BVIT-dependent, such that BO4502 fusions generated the highest mIL-10 levels and BT1491, the highest mIL-22. To verify bioactivity of mIL-22, we assayed STAT3 phosphorylation in human epithelial HT-29 cells, as IL-22 signaling leads to STAT3 activation [287]. In control experiments (Fig. 4-9C), we could readily detect STAT3 phosphorylation when cells were stimulated with recombinant IL-22. However, mIL-22 derived from *Bacteroides* culture supernatants (~5-10ng/mL) did not produce any detectable transcription factor phosphorylation (Fig. 4-9D). To verify if spent *Bacteroides* culture media contained factors that inhibit IL-22 signaling, we assayed activity of recombinant IL-22 mixed with culture supernatants, which yielded STAT3 phosphorylation. Similar STAT3 phosphorylation assays were conducted with mIL-10 containing culture supernatants with J774 murine macrophages with similar negative results (data not shown). While BVIT fusions of mammalian cytokines were able to lead to enrichment of cytokine in culture supernatants, produced protein was not bioactive.

The lack of bioactivity could be attributable to a number of factors. Firstly, mammalian cytokines, such as mIL-10 and mIL-22, are notoriously difficult to produce in Gram negative bacteria as their synthesis leads to accumulation with inclusion bodies [**Sorensen2005**]. OMVs are a mechanism employed by bacteria to expel misfolded protein [265], therefore the protein detected by ELISA could be discarded misfolded cytokine. Moreover, compared to facultative anaerobic bacteria, like *E. coli*, both the cytoplasm and periplasmic compartments of *Bacteroides* are reduced environments, which is not conducive to the synthesis of proteins containing disulfide bonds, such as mIL-10 and mIL-22 [288]. Multiple troubleshooting strategies were attempted to rescue bioactivity of expressed mIL-10 and mIL-22, including: varying the linker sequence between BVITs and the cytokine [289]; importing the DsbABC disulfide bond chaperone system from E. coli [290]; varying media and growth conditions; testing

activity of purified OMVs instead of culture supernatants; and, oxidatively stressing cells to promote disulfide bond formation. None of these strategies yielded bioactive protein. *Bacteroides* spp. or other obligate anaerobes may not useful for *in situ* protein delivery of mammalian proteins.

4.5 *Bacteroides* OMVs for Antigen-Specific Tolerance

Carriers of many toll-like receptor agonists, OMVs are intrinsically potent inducers of immune responses. Indeed, OMVs derived from pathogenic Enterobacteriaceae have been explored as adjuvants for vaccines to packaged cargo [271, 291]. However, *Bacteroides* OMVs are unique in their ability to dampen inflammatory responses compared to OMVs derived from other species [268]. The ubiquity of *Bacteroides* as commensal organisms suggests that the host develops immunological tolerance to *Bacteroides* antigens to maintain healthy gut homeostasis. The development of tolerance to oral dietary antigens is well described and dependent on dendritic cell antigen presentation to naïve T cells that differentiate into FoxP3+ IL-10+ CD4+ Treg cells [292–294]. OMVs derived from *B. fragilis* can similarly induce Treg differentiation in a dendritic cell dependent manner [268, 269]. By packaging model antigens into OMVs, we can explore how *Bacteroides*-derived OMVs can influence antigen presentation and downstream immune responses.

As a proof-of-concept, we fused the model antigen ovalbumin (Ova) to the BVIT BT1491 and examined expression in the cellular, supernatant, and purified OMV fractions when expressed from *B. thetaiotaomicron*, *B. fragilis*, and *B. vulgatus* (Fig. 4-10). Ova expression could be detected in all fractions, although the protein was degraded in culture supernatants. Interestingly, Ova seemed to be protected from degradation in purified OMVs, suggesting that cargo may be cleaved and liberated from the vesicles in culture supernatants. As a control, we also generated strains expressed Ova intracellularly (inOVA), which could only be detected in the pellet

fraction.
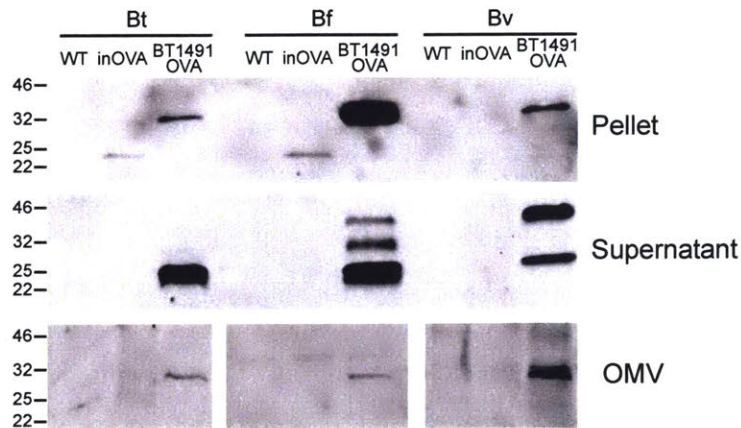
Figure 4-10: **Incorporation of Ova into *Bacteroides* OMVs:** The model antigen ovalbumin (Ova) was expressed intracellularly (inOVA) or fused to the BVIT BT1491d32. Expression in cell pellets, culture supernatants, and purified OMVs derived was analyzed by Western blot against the C-terminal FLAG tag on Ova. 5 μg of *B. thetaiotaomicron* and 1μg of *B. fragilis* and *B. vulgatus* OMVs were analyzed. Blots are representative of at least three separate experiments.
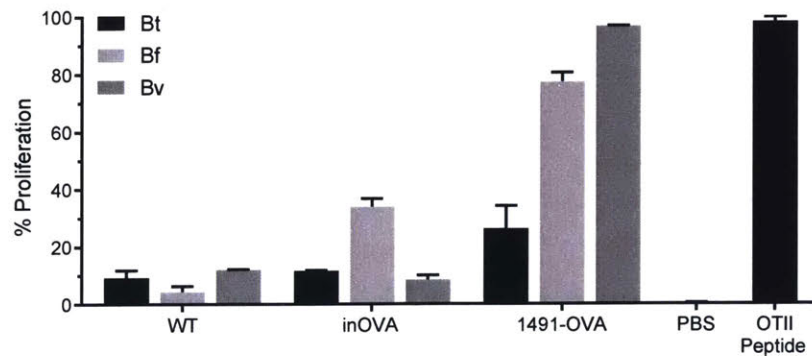
Figure 4-11: **Presentation of Antigen-Loaded OMVs to Naïve T Cells:** Dendritic cells were isolated from WT C57BL/6 mice previously injected with B16-FLT3L cells. Dendritic cells were stimulated with purified OT-II peptide or with 1 μg OMV. Naïve CD4+ T cells isolated from OT-II mice were labeled with CellTrace Violet and added to cultures 2 hours dendritic cell stimulation. Proliferation was analyzed by flow cytometry 3.5 days following addition of T cells. Error bars represented the standard deviation of 2-4 independent biological replicates.

Next, we evaluated if antigen packaged into *Bacteroides* OMVs could be properly processed and presented by dendritic cells. Dendritic cells were incubated with *Bac-*

teroides OMVs and subsequently co-cultured with naïve CD4+ T cells isolated from transgenic OT-II mice, whose T cell receptor repertoire is restricted to an epitope of Ova. Wild-type OMVs and those derived from strains expressing Ova intracellularly failed to induce significant T cell proliferation. Some proliferation was observed at with Bf inOVA, suggesting that some cytoplasmic antigen may be co-purified with OMVs. Incorporation of Ova into OMVs led to robust proliferation, comparable to that of purified OT-II peptide for Bf and Bv OMVs (Fig. 4-11). The resultant cytokine profile of co-cultures was analyzed to discern the helper T cell subtype. Wild-type OMVs induced IL-10 secretion compared to untreated and peptide treated controls and the presence of antigen further increased IL-10 levels (Fig. 4-12A). In contrast, IFNγ, an inflammatory cytokine, was induced upon activation, but at significantly lower levels in OMV samples compared to peptide control (Fig. 4-12B). IL-17, a Th17 cytokine, was also reduced compared to peptide control and only induced at a low level in the presence of OMVs (Fig. 4-12C). Similarly, the Th2 cytokine IL-4 was not highly induced upon incubation with OMVs (Fig. 4-12D). Together, activated T cells highly express IL-10 and show reduced secretion of IFNγ, suggesting a regulatory-like phenotype.

4.6 Discussion

Here, we developed a strategy to direct heterologous protein packaging into OMVs derived from *Bacteroides* spp. Using a proteomics and bioinformatics approach, we found that *Bacteroides* OMVs were enriched in SpII-type lipoproteins. Translational fusion to these lipoproteins was sufficient to package heterologous cargo to OMVs. Moreover, sequential truncation of OMV lipoproteins led to the identification of 25-35 amino acid BVITs that could confer OMV incorporation at similar or superior levels compared to full length OMV proteins. Using design rules informed from empirical and predicted sequences, we demonstrated that BVITs derived from *B. thetaiotaomicron*, *B. fragilis*, and *B. ovatus* lipoproteins could direct OMV incorporation in several species in the *Bacteroides* genus. Next, we employed BVITs to package mammalian
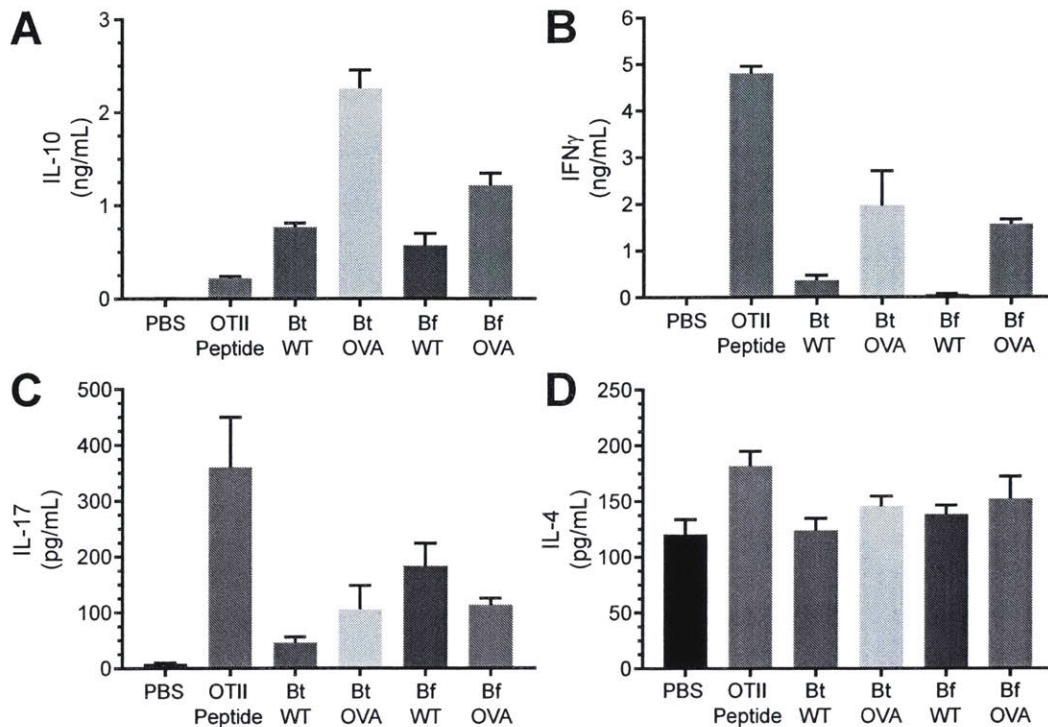
Figure 4-12: **Cytokine Profiling of Dendritic Cell-T Cell Co-Culture Assays:** Levels of IL-10 (A), IFNγ (B), IL-17 (C), and IL-4 (D) were analyzed by ELISA. Error bars represent standard deviation of three biological replicates (n=3).

cytokines into OMVs. Although BVITs could successfully direct packaging, *Bacteroides*-produced cytokines were not biologically active, suggesting that *Bacteroides* may be poor chassis for expression of mammalian proteins. Finally, we used BVITs to direct the incorporation of a model antigen into OMVs to explore how the host immune system responds to OMV antigens. In an antigen presentation assay, we showed that antigen-loaded OMVs could be properly processed and presented by dendritic to induce antigen-specific proliferation of naïve CD4+ T cells. Cytokine profiling activated T cells suggested a regulatory-like phenotype.

The ability to direct proteins in commensal OMVs could be a versatile platform to modulate host immune responses. Given their ubiquity, abundance and stability [16] in human gut microbiomes, *Bacteroides* spp. OMVs could be potent agent to affect the immune system. Although incorporation of mammalian cytokines into OMVs proved unsuccessful, many other cargo could be explored as interesting immunomod-

ulators. For example, intestinal alkaline phosphatase has been shown to detoxify lipopolysaccharide and dephosphorylate extracellular adenosine tri-phosphate [295] and oral administration has been protective in mouse models of colitis [296]. As public goods, OMVs could be loaded with enzymes that could benefit the microbial community or to detoxify compounds harmful to the host [119].

Engineered OMVs could also be useful tools to target the adaptive immune system, as the native components of OMVs are highly immunostimulatory. By packaging model antigens into commensal OMVs, we can explore how commensal antigens are treated and eventually tolerated by the host immune system. If *Bacteroides* OMVs can induce antigen-specific tolerance, they be applied therapeutically for allergic or autoimmune diseases. Antigen-loaded nanoparticles [297] and erythrocytes [298], as well as sialic acid modification of antigens [299], has been explored to induce tolerance in autoimmune models. Commensal OMVs could complement these efforts and, due to the proximity to the gut mucosa, could be more effective for immune diseases of the gastrointestinal tract. Our efforts to package ovalbumin into OMVs sets the stage to investigate their toleregenic potential and we will further evaluate their function in mouse models of antigen-induced allergic diarrhea [300]. Together, engineered OMVs could serve as therapeutic agents for immune diseases of the gut as well as tools to study the immune responses to commensal antigens.

4.7 Experimental Details

Bacterial Strains and Culture Conditions

B. thetaiotaomicron VPI-5482, *B. fragilis* NCTC9343 and *B. vulgatus* ATCC8482 were routinely cultured in TYG media supplemented with 0.4% glucose in anaerobic conditions at 37°C. (see Chapter 3 Experimental Details: Bacterial Strains and Culture Conditions). All *Bacteroides* media was pre-reduced overnight in the anaerobic chamber overnight before use. For expression studies, TYG was supplemented with 0.2 mM IPTG. *E. coli* S17-1 λpir as routinely cultured in LB broth (Difco)

supplemented with 100 μg carbenicillin where appropriate.

Outer Membrane Vesicle Preparations

Outer membrane vesicles were purified as previously described (Chutkan et al., 2013). Overnight cultures of *Bacteroides* spp. were subcultured 1:100 in 80mL of pre-reduced TYG in 100mL glass bottles and grown for 16-20h anaerobically at 37°C. Cultures were centrifuged at 10,000 x g for 15 minutes to pellet bacterial cells. Supernatants were filtered through 0.45 μm PVDF membranes (Millipore) to remove residual cells and cellular debris. The filtrate was subsequently ultracentrifuged at 100,000 x g for 2 hours in a 70Ti rotor (Beckman Coulter) to pellet OMVs. The pellets were washed once in equal volumes of cold PBS. The final pellets were resuspended in 0.5mL of sterile PBS. OMV yield was quantified by total protein content as measured by Bradford assay (BioRad).

Mass Spectroscopy Analysis

Purified OMVs were submitted for LC-MS/MS by the Koch Institute Biopolymers & Proteomics Core. The procedure was performed on three preparations of OMVs isolated from separate cultures. Samples were reduced with 10mM dithiothreitol (Sigma) for 1h at 56°C and then alkylated with 55mM iodoacetamide (Sigma) for 1h at 25°C in the dark. Proteins were then digested with modified trypsin (Promega) at an enzyme/substrate ratio of 1:50 in 100mM ammonium acetate, pH 8.9 at 25°C overnight. Trypsin activity was halted by addition of acetic acid (99.9%, Sigma) to a final concentration of 5%. Peptides were desalted using C18 SpinTips (Protea, Morgantown, WV) then vacuum centrifuged and stored at -80 °C.

Peptides were separated by reverse phase HPLC (Thermo Easy nLC1000) using a precolumn (made in house, 6 cm of 10 μm C18) and a self-pack 5 μm tip analytical column (12 cm of 5 μm C18, New Objective) over a 140 minute gradient before nanoelectrospray using a QExactive mass spectrometer (Thermo). Solvent A was 0.1% formic acid and solvent B was 80% MeCN/0.1% formic acid. The gradient

conditions were 2-10% B (0-3 min), 10-30% B (3-107 min), 30-40% B (107-121 min), 40-60% B (121-126 min), 60-100% B (126-127 min), 100% B (127-137 min), 100-0% B (137-138 min), 0% B (138-140 min), and the mass spectrometer was operated in a data-dependent mode. The parameters for the full scan MS were: resolution of 70,000 across 350-2000 m/z, AGC 3e6, and maximum IT 50 ms. The full MS scan was followed by MS/MS for the top 10 precursor ions in each cycle with a NCE of 28 and dynamic exclusion of 30 s. Raw mass spectral data files (.raw) were searched using Proteome Discoverer (Thermo) and Mascot version 2.4.1 (Matrix Science). Mascot search parameters were: 10 ppm mass tolerance for precursor ions; 15 mmu for fragment ion mass tolerance; 2 missed cleavages of trypsin; fixed modification was carbamidomethylation of cysteine; variable modifications were methionine oxidation. Only peptides with a Mascot score greater than or equal to 25 and an isolation interference less than or equal to 30 were included in the data analysis.

Identified peptide sequences were mapped to the *B. thetaiotaomicron* reference genome to discern the proteins present in OMV fractions. Proteins identified in all samples that were mapped with at least 18 total peptides and 3 unique peptides were considered to be OMV proteins.

Lipoprotein Bioinformatic Analysis

Sequence analysis was performed in Geneious 8.1 and Matlab R2017a. Coding regions from the *B. thetaiotaomicron* and *B. fragilis* reference genomes were extracted and translated into corresponding amino acid sequences. Protein sequences were analyzed by LipoP1.0 and SignalP2.0 to identify the presence of SpI, SpII or transmembrane domains. Proteins with computationally-predicted SpII cleavage sites were considered as *Bacteroides* lipoproteins (n=1069). Proteins identified in mass spectrometry analysis as well as those identified in a previous proteomics screen [275] were labeled as OMV lipoproteins (n=100) and the remaining lipoproteins were considered to be cellular lipoproteins. The amino acid frequency at residues relative to the conserved cysteine residue were calculated in each group. The weight difference

in amino acid frequency at a given position was determined by the following formula:

$$WeightedDifference = F_{OMV} \left(log_2(\frac{F_{OMV}}{F_{Cell}}) \right)$$

Where F_{OMV} is the amino acid frequency at a given position of OMV lipoproteins and F_{Cell} is the amino acid frequency at a given position of cellular lipoproteins. Sequence logos were constructed using WebLogo (https://weblogo.berkeley.edu/logo.cgi).

Genetic Parts and Plasmids

All plasmids were constructed using PCR, DpnI digest, Gibson assembly [148], and transformation into *E. coli* S17 λpir (as described in Chapter 3, Experimental Details). BVITs were amplified from purified *B. thetaiotaomicron* or *B. fragilis* genomes using Kapa Hifi Polymerase. Genes encoding mIL-10, mIL-22 and OVA were synthesized by Integrated DNA Technologies as Geneblocks. Conjugation of constructs in *Bacteroides* spp. were performed as described in Chapter 3, Experimental Details. All protein coding genes were cloned into an IPTG-inducible pNBU1 expression vector.

Luminescence Assays

Overnight cultures of *Bacteroides* strains were subcultured 1:100 in pre-reduced TYG media and grown at 37 °C until an OD_{600} of $\sim$0.6-0.8 ($\sim$6h). Samples were centrifuged at 5,000 x g for 5 minutes. Supernatant was removed and saved for later analysis. In some cases, supernatant samples were concentrated using ultrafiltration columns (Amicon, MWCO 10kDa) as per manufacturer's instructions. Pellets were washed twice in PBS and finally resuspended in an equal volume of PBS.

Supernatants, pellets and OMVs were assayed for luminescence activity using the Nano-Glo Luciferase Assay System (Promega) as per manufacturer's instructions. Samples were not lysed before assaying, as the NanoLuc enzyme buffer contains detergents.

Proteinase K Digestion

Suspensions of 5 μg of OMVs in PBS were treated with 0.1 mg/mL proteinase K for various times between 9 h at 37°C in the presence or absence of 1% Triton-X100. Following the incubation, all samples were placed on ice and proteolysis was stopped by addition of 1mM phenylmethanesulfonyl fluoride (PMSF). The effects of proteinase K and detergents treatments were determined by luciferase assay, as above.

Western Blots

Protein samples were analyzed by Western blot using the Bio-Rad Mini-PROTEAN System. *Bacteroides* cell pellets (equivalent of 0.075 OD_{600} units), *Bacteroides* culture supernatants (equivalent of 0.5mL concentrated using Amicon columns), HT-29 cell lysates (10 μg), or purified OMVs (1-5 μg) were mixed with 4x Laemmli Sample Buffer (Bio-Rad) with 10% β-mecaptoethanol and boiled for 10 minutes. Protein samples were separated on 4-15% Mini-PROTEAN TGX gels in a Tris-SDS-glycine buffer system for 30 minutes at 200V. Proteins were transferred to nitrocellulose membrane using a Tris-glycine-methanol buffer for 60 minutes at 60V. Samples were blocked overnight in 5% milk in PBS with 0.05% Tween 20 (PBST) for anti-His blots, overnight in 2.5% milk in PBST for anti-FLAG blots, and for 2 hours in 10% BSA Blocking Solution (SeraCare) for anti-P-STAT3 blots. Primary antibody incubations were conducted in appropriate blocking reagent as follows; 1:5000 dilution of Ms anti-His (HIS.H8, Abcam ab18184) for 4 hours at room temperature; 1:2500 dilution of Ms anti-FLAG (M2, Sigma F3165) for 4 hours at room temperature; 1:2000 Rb anti-Phospho-Stat3 (Tyr705) XP (Cell Signaling Technologies, # 9145) overnight at 4°C. Secondary antibody incubations were conducted in blocking reagent at room temperature as follows: 1:5000 Rabbit anti-Ms IgG-HRP (Abcam, ab6728) for 1 hour; and, 1:5000 Anti-Rb IgG-HRP (Cell Signaling Technologies, # 7074) for 1 hour. Following primary and secondary antibody incubations, blots were washed four times in PBST for at least 5 minutes per wash. Blots were developed using SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Scientific)

for 5 minutes and images were acquired using a Bio-Rad Gel Doc XR+ Gel Documentation System. Image analysis and annotation was performed using ImageLab (Bio-Rad) and Adobe Illustrator for layout.

IL-22 Bioactivity Assay

HT-29 cells were routinely cultured in Dulbecco's Modified Eagle Medium (DMEM; Gibco) supplemented with 10% fetal bovine serum (Gibco) and Pen-Strep (Gibco) in 10 cm plates in a 5% CO2 incubator at 37°C. One day prior to the assay, 90% confluent monolayers were trypsinized and seeded in 6-well plates at a concentration of $5x10^5$ cells per well. The following morning, cells were washed with PBS and cultured for 4-6 hours in DMEM without serum supplementation. To prepare *Bacteroides* culture supernatant, strains were cultured for 6 hours at 37°C and centrifuged to isolate supernatant. 0.5mL of culture supernatant (equivalent of $\sim$5-10 ng of mIL-22 as detected by ELISA) was concentrated to $\sim$50 μL using ultrafiltration columns. HT-29 cells were then treated with serum-free DMEM supplemented with recombinant murine IL-22 (Peprotech) or *Bacteroides* culture supernatant for 15 minutes at 37°C. Monolayers were placed on ice, washed with cold PBS and lysed in RIPA buffer containing PMSF (Sigma), cOmplete protease inhibitor cocktail (Roche), and HALT phosphatase inhibitor cocktail (Thermo Fisher) for 20 minutes on ice. Cells were scraped into microcentrifuge tubes and subsequently centrifuged at 15,000 x g for 10 minutes to pellet cell debris. The supernatant was transferred to a new tube and protein concentration was measured using a BCA assay (Thermo Fisher).

Antigen Presentation Assays

Dendritic cells were purified with the PanDendritic Cell Isolation Kit (Miltenyi Biotec) from spleens and mesenteric lymph nodes of WT C57BL/6 mice injected with $1x10^6$ B16-FLT3L cells. T cells were purified with the Naïve CD4+ T Cell Isolation Kit (Miltenyi Biotec) from the spleens and mesenteric lymph nodes of OT-II mice and labeled with CellTrace Violet (Thermo Fisher). Purity was always greater than

98% for both dendritic cells and T cells. Dendritic cells ($1\mathrm{x}10^5$) were incubated with purified OMVs (1 μg/ml), or 100nM OT-II peptide for 2 hours before the addition of $1\mathrm{x}10^5$ CD4+ T cells. Cell division was assessed by violet trace dilution and analyzed by flow cytometry after 3.5 days of co-culture. Cytokine profiles of culture supernatants were determined by ELISA.

ELISA Assays

Concentrations of murine IL-4, IL-10, IL-17, IL-22, IFNγ, and TGFβ were measured as per manufacturer's instructions (R & D Systems, DuoSet ELISA).

Chapter 5

Bacterial-Electronic Device to Monitor Gastrointestinal Health

The following chapter is adapted from the 2018 manuscript to which Phillip Nadeau and I contributed equally as first authors.

Abstract

Biomolecular monitoring in the gastrointestinal tract could offer rapid, precise disease detection and management, but is impeded by access to the remote and complex environment. Here, we present an Ingestible Micro-Bio-Electronic Device (IMBED) for *in situ* biomolecular detection based on environmentally-resilient biosensor bacteria and miniaturized luminescence readout electronics that wirelessly communicate with an external device. As a proof-of-concept, we engineer heme-sensitive probiotic biosensors and demonstrate accurate diagnosis of gastrointestinal bleeding in swine in as little as 52 minutes. Additionally, we integrate alternative biosensors to demonstrate modularity and extensibility of the detection platform. IMBEDs enable new opportunities for gastrointestinal biomarker discovery and could transform the management and diagnosis of gastrointestinal disease.

5.1 Introduction

Microorganisms living on and in the human body constantly interrogate their biochemical surroundings and alter gene expression to adapt to changing environments. Synthetic biology enables the robust engineering of living cells with increasingly complex genetic circuits to sense multiple biological inputs and control gene expression [205]. Whole-cell biosensors harness this sensing ability to detect analytes associated with human health [301] or environmental contamination [302]. Due to their innate robust functionality in complex physiological environments, biosensors have been developed to sense clinically relevant biomarkers in serum or urine *ex vivo* [108] as well as gut biomolecules supplemented in diet [109, 196, 303] or generated during disease [110, 111, 304]. However, despite their promise as non-invasive diagnostics, biosensors have yet to be employed for clinically compatible testing in an unobtrusive, real-time, and user-friendly way. Current research applications of ingestible biosensors in animal models rely on cumbersome analysis of bacterial gene expression or DNA in stool samples [109–111, 196, 303, 304], rather than real-time reporting from within the body. Furthermore, the scalability of simultaneous detection of multiple biomarkers is limited by the bounds of genetic circuit complexity, cellular burden of synthetic constructs, and orthogonal gene expression outputs [305].

In contrast, the impressive scaling of semiconductor microelectronics over the past few decades has delivered sophisticated, highly miniaturized platforms for ultra-low power sensing, computation, and wireless communication [306–308]. In the gastrointestinal tract, electronic capsules have been deployed for taking visual images [309], delivering drugs while measuring temperature and pH [310], sensing intestinal gases [311] and recording patient compliance [312]. While electronics provide a versatile interface for collecting, processing, and sharing information, their ability to directly sense biomolecules *in vivo* has been limited due to their dependence on labile biochemical transducers that necessitate large, power-demanding circuits for sensitive detection.

Here, we describe an Ingestible Micro-Bio-Electronic Device (IMBED) that com-

bines probiotic bacteria engineered with biosensor genetic circuits together with ultra-low power microelectronics to enable *in situ* detection of gastrointestinal biomolecules associated with health or disease (Fig. 5-1). By partitioning sensing to biological systems, and computation and communication to electrical devices, IMBEDs leverage the natural advantages of each approach to enable a novel class of ingestible gastrointestinal diagnostics. In our IMBED, biosensor probiotics lie adjacent to readout electronics in individual wells separated from the outside environment by a semi-permeable membrane that confines cells in the device and allows for diffusion of small molecules. Sensing of target biomarkers by the bacteria generates light, which is then detected by photodetectors embedded in the electronics. These electrical signals are processed by a integrated bioluminescence detection circuit [313] and are transmitted wirelessly from the device to an external radio or cellular phone for convenient readout.
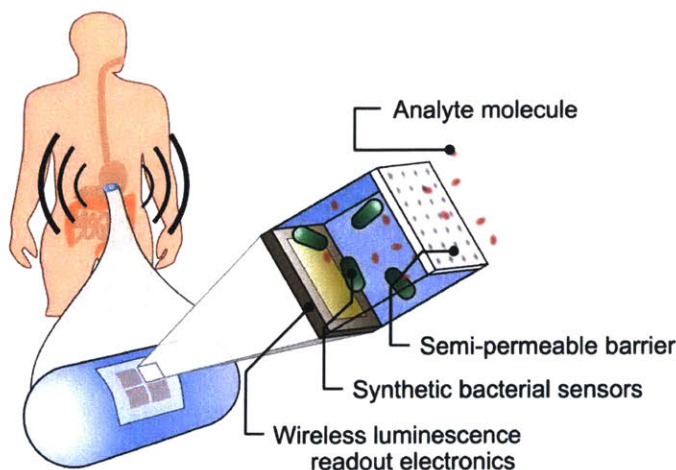
Figure 5-1: Capsule for sensing biomarkers *in vivo* with whole-cell bacterial sensors and wireless electronic readout.

5.2   Blood Biosensor Design and Validation

We developed a biosensor for gastrointestinal bleeding as a proof-of-concept IMBED for a clinically relevant biomarker. Bleeding in the gastrointestinal tract can be a result of a wide range of causes, including inflammation, cancer, peptic ulcers,

non-steroidal anti-inflammatory drug use, portal vein hypertension, among others [314]. While cost-effective fecal occult-blood testing exists [315], rapid diagnosis of acute bleeding in the upper gastrointestinal tract requires endoscopic observation or aspiration of gastric fluid [316]. Importantly, early diagnosis and appropriate treatment of individuals with upper gastrointestinal bleeding has been found to reduce hospital stays and overall medical costs [317]. Blood sensing IMBEDs could offer an additional means of diagnosing upper gastrointestinal bleeds or monitoring patients at high risk for re-bleeding following endoscopic therapy [318] to aid in the triage of individuals who may require further endoscopic or surgical intervention.

5.2.1   Gene Circuit Design
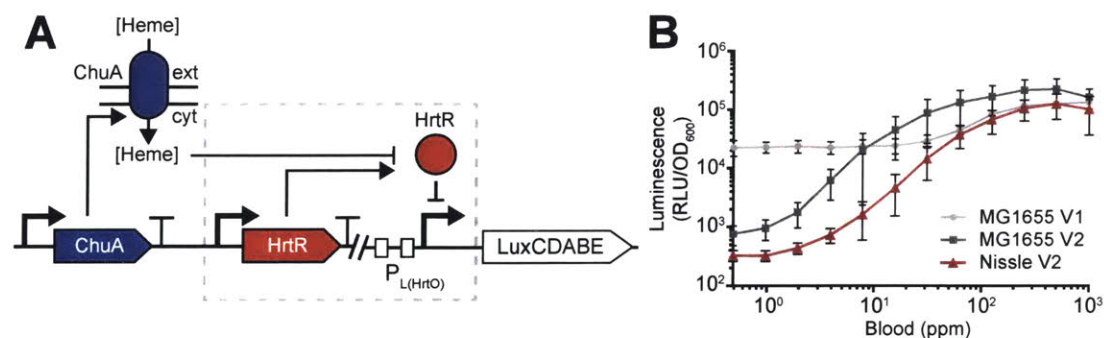
Figure 5-2: **Probiotic *E. coli* can be engineered to sense blood *in vitro*:** (A) Schematic of the blood sensor gene circuit. Extracellular heme is internalized through the outer membrane transporter ChuA and interacts with the transcriptional repressor HtrR to allow for transcription of the bacterial luciferase operon *luxCDABE*. (B) Dose-response curves of prototype (V1) and optimized (V2) heme sensing genetic circuits in laboratory (MG1655) and probiotic (Nissle) strains of *E. coli*. Error bars represent SEM of three independent biological replicates.

As a bleeding event leads to an accumulation of free heme liberated from lysed red blood cells, we examined the literature for bacterial transcription factors responsive to heme. *Lactococcus lactis* encodes a heme-regulated TetR-family transcriptional repressor, HrtR, which naturally controls expression of an efflux pump to control intracellular heme-mediated toxicity [319]. In the absence of heme, HrtR binds to cognate HrtO operator sequences in the P_{hrtRAB} promoter, repressing promoter activity (Fig.

5-2A). Conformational changes in HrtR upon heme binding abrogate DNA binding and lead to downstream gene expression [320]. To adapt the native P_{hrtRAB} promoter to an *Escherichia coli* chassis, we created a synthetic promoter, $P_{L(HrtO)}$, based on the strong late promoter of bacteriophage lambda with HrtO operator sequences directly upstream of the -35 and -10 boxes (Fig. 5-3A). Although photon flux is lower than eukaryotic luciferases, the *Photorhabdus luminescens luxCDABE* luciferase operon was used as the output of $P_{L(HrtO)}$ as it functions at body temperature and encodes all components necessary for intracellular substrate production, thus obviating the need for exogenous substrate [321]. Co-transformation of $P_{L(HrtO)}$-*luxCDABE* with a constitutively expressing HrtR construct in *E. coli* MG1655 led to a 4.4-fold reduction in luminescence, indicating HrtR-mediated repression of $P_{L(HrtO)}$ (Fig. 5-3B). However, luminescence levels remained constant irrespective of heme concentration, suggesting that heme could not penetrate the Gram-negative cell envelope. Pathogenic strains of *E. coli* have evolved heme scavenging systems to acquire scarcely available iron during infection [322]. We hypothesized that introducing the ChuA heme transporter from *E. coli* O157:H7 into the gene circuit would allow for the transit of extracellular heme into the periplasm, where it could subsequently interact with other cellular components to enter the cytoplasm and finally complex with HrtR (Fig. 5-2A) [323]. Expression of both HrtR and ChuA yielded a biosensor (MG1655 V1) that responded to increasing heme input with luminescence output with a signal-to-noise ratio (SNR) of 5.9 and a K_D of 1μM heme (Fig. 5-3B). Luminescence production was also induced by whole horse blood (Fig. 5-3C) and lysis of red blood cells in simulated gastric fluid greatly improved the sensitivity of detection by liberating heme (K_D=115ppm blood) (Fig. 5-2B; Fig. 5-3D).

Next, the prototype genetic circuit was iteratively optimized with the goal of improving SNR without compromising maximum luminescence output. Genetic components were combined onto a single high-copy plasmid to minimize plasmid burden as well as the risk of plasmid loss. Increasing the translation initiation strength of HrtR using computationally designed ribosome binding site (RBS) sequences [241] decreased baseline luminescence and improved SNR to 132 (MG1655 V2; Fig. 5-2B;
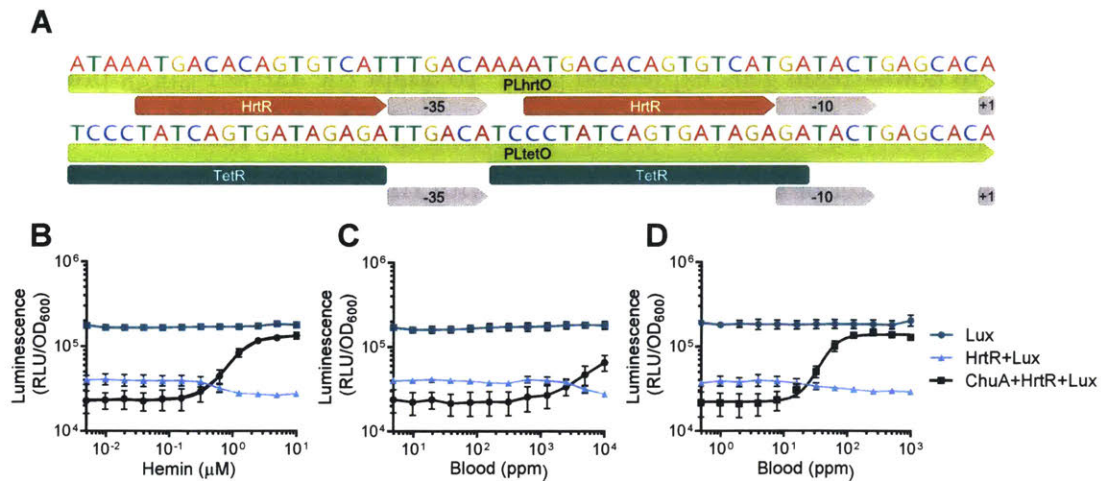
Figure 5-3: **Design and *in vitro* evaluation of prototype heme sensing genetic circuit:** (A) Promoter design of heme-responsive promoter. The TetR operator sites of a synthetic promoter based on the left promoter of bacteriophage lambda ($P_{L(TetO)}$) [160] were replaced with the operator DNA sequences to which HrtR binds. Spacing between the -10 and -35 sites was preserved. (B-D) Dose-response curves of prototype genetic circuits in *E. coli* MG1655 in various concentrations of hemin (B), whole horse blood (C), and blood lysed in simulated gastric fluid (D). The genetic circuit contains $P_{L(HrtO)}$-*luxCDABE* alone (Lux), $P_{L(HrtO)}$-*luxCDABE* with the HrtR transcriptional repressor (HrtR+Lux), or $P_{L(HrtO)}$-*luxCDABE*, HrtR and the ChuA heme transporter (ChuA+HrtR+Lux). Luminescence values are measured 2 hours post-exposure to inducer and normalized to the optical density of the culture. Error bars represent SEM of three independent biological replicates.

Fig. 5-4A-D). Variations in promoter sequence, number and position of HrtO operator sites in $P_{L(HrtO)}$, as well as ChuA RBS strength did not lead to appreciable improvements in gene circuit performance. The final gene circuit was transferred to the probiotic *E. coli* Nissle 1917 strain (Nissle V2) and retained similar performance characteristics compared to the laboratory strain in response to lysed horse blood (SNR=310; K_D=95ppm) (Fig. 5-2B) as well as human blood (Fig. 5-4F). Luminescence was induced rapidly, reaching half-maximal levels within 45 minutes of exposure to heme or lysed blood (Fig. 5-4E).
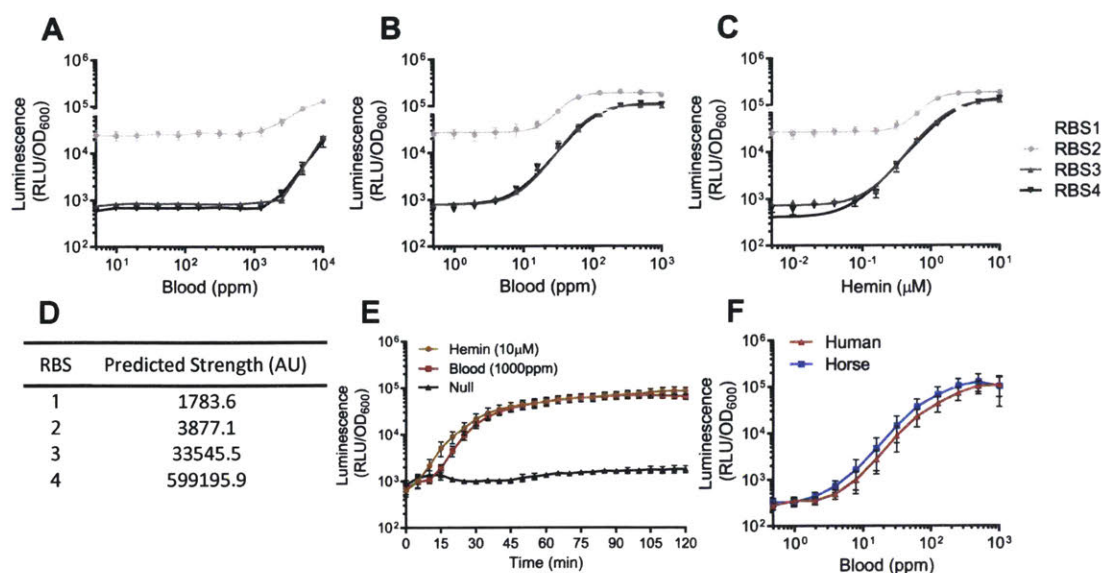
Figure 5-4: **Optimization and Characterization of Heme-Sensing Gene Circuit:** Dose-response curves of heme-sensing genetic circuits in *E. coli* MG1655 in various concentrations of hemin (A), whole horse blood (B), and blood lysed in simulated gastric fluid (C). The translational initiation strength of HrtR was varied using different computationally-designed ribosome binding sites (RBS) [241, 324]. Predicted RBS strengths are listed in D. Luminescence values are measured 2 hours post-exposure to inducer and normalized to the optical density of the culture. (E) *E. coli* Nissle blood biosensors (Nissle V2 from Fig. 5-2B) were treated with 10 μM hemin (brown), 1000 ppm blood (red) or PBS (black) and luminescence response was measured in a plate reader every 5 minutes for 2 hours. Luminescence values are normalized to the optical density of the bacterial culture. (F) *E. coli* Nissle blood sensor strains (Nissle V2 from Figure 1B) were treated with various concentrations of human or horse blood lysed in simulated gastric fluid. Luminescence values are measured 2 hours post-exposure to inducer and normalized to the optical density of the culture. Error bars represent SEM of three independent biological replicates.

5.2.2   Blood Sensor Performance in Mice

To examine functionality of the bacterial blood sensor *in vivo*, we employed a murine model of indomethacin-induced gastrointestinal bleeding. Gastroduodenal ulceration is a common adverse effect of non-steroidal anti-inflammatory drug use, as decreased prostaglandin production leads to a thinning of the gastric mucosa and acidification of gastric contents [325]. Upper gastrointestinal bleeding elicited by oral indomethacin administration could be detected by bacterial blood sensors passing through the gut and measured by observing luminescence activity in fecal pellets

(Fig. 1C). Bacterial transit to stool was maximal 6 hours post-inoculation and the blood sensor bacteria could not be recovered from mouse stool 24 hours after administration, suggesting that the engineered strains did not appreciably colonize the murine gut (fig. S6). At baseline, administration of blood sensor bacteria did not lead to detectable luminescence activity in stool, indicating that the basal heme levels in the murine gut are insufficient to activate the gene circuit (fig. S7). Oral administration of indomethacin generated overt gastrointestinal bleeding overnight as noted by black, tarry stool and positive guaiac tests. Mice subsequently inoculated with blood sensor bacteria demonstrated 18-fold higher luminescence values in fecal pellets as compared to controls (Fig. 1C). Biosensor detection events were fully concordant with guaiac tests and could perfectly discriminate between indomethacin treated and untreated animals. Our designed biosensor can thus effectively detect the presence of gastrointestinal bleeding *in vivo*.

5.3   Integration with Electronic Readout Capsule

5.3.1   Electronic Capsule Design

We next sought to integrate our bacterial biosensor with an electronic sensor and wireless transmission platform. Interrogation of cellular bioluminescence is typically performed by power and area-expensive lab equipment that is ill-suited for *in situ* measurements in the body. Prior demonstrations of custom sensitive bioluminescence detection electronics have required external wiring and have been limited to bench-top assays [313, 326, 327]. For this reason, we developed a miniaturized, fully-integrated, wireless readout capsule for targeted *in vivo* sensing of small molecules in the gastrointestinal tract (Fig. 5-6). The system encapsulates our prior nanowatt-level time-based luminometer [313], with a microprocessor and wireless transmitter, and provides containment for engineered cells for molecular sensing. Our IMBED consists of two parts: (1) a molded capsule containing the electronic components, and (2) a plastic carrier for containing cells in one of four cavities. Bioluminescence
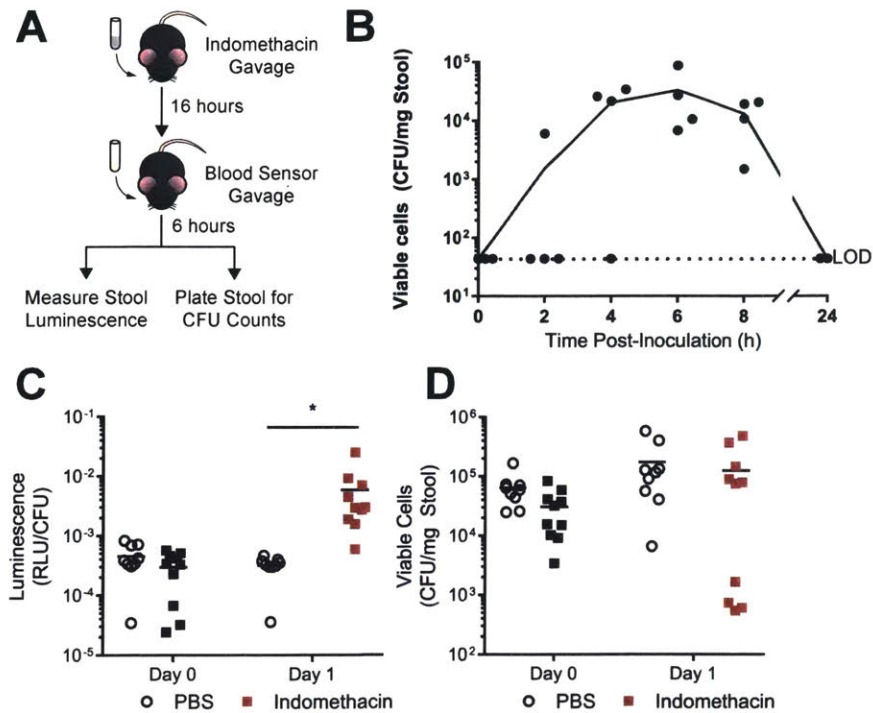
Figure 5-5: **Heme biosensors can detect blood in an *in vivo* murine model of indomethacin-induced gastrointestinal bleeding:** (A) Experimental design. (B) C57BL/6J mice were inoculated with approximately 2×10^8 CFU of blood biosensors by oral gavage (n=4). Fecal pellets were collected from mice prior to gavage and at 2, 4, 6, 8 and 24 hours post-gavage and plated to determine CFU counts. All mice contained biosensor bacteria in their stool 6 hours post-gavage and no colonization was observed. Dotted line indicates the limit of detection (LOD) of the assay. (C) Mice were inoculated with approximately 2×10^8 CFU of E. coli Nissle blood sensors 6 hours prior to (Day 0) or 16 hours after (Day 1) administration of indomethacin (10 mg/kg) or PBS buffer as a negative control. Induction of bleeding was confirmed by guaiac test (indicated by red color). Fecal pellets were collected from animals 6 hours post-gavage, homogenized and analyzed for luminescence production as well as plated to enumerate colony forming units (CFU). Luminescence values were normalized to cell number in fecal pellets (N = 10). *P<0.05, Student's t-test. (D) CFU counts in fecal pellets 6 hours post-gavage.

from activated cells is detected by phototransistors located below each cavity and converted to a digital code using the low-power luminometer chip. In each IMBED, one channel acts as a reference to calibrate for background light and temperature-induced dark current variation, while the remaining three are used for independent measurements. Incident photocurrent is supplied to an on-board microcontroller and

900 MHz wireless radio for transmission outside the body. A small button-cell battery (5 mAh) powers the device and the extrapolated IMBED power consumption suggest a nominal device shelf-life of over 9 months and active operation time of 1.5 months on a full charge. The low power consumption achieved also could allow for battery-free operation in the gastrointestinal tract using energy harvested from gastric acid [328]. In addition, two 220 μF ceramic capacitors supplied the instantaneous peak energy required by the radio transmitter. Electronic components were coated in Parylene-C to provide necessary humidity resilience for the sensitive picoampere-level photocurrent measurements. Devices were subsequently encapsulated with a rigid epoxy for mechanical robustness, followed by a molded PMDS capsule for biological compatibility. This multi-layered electronics packaging strategy allows for the creation of a robust cm-scale wireless capsule that, when paired with biosensor cells, can perform continuous, minimally-invasive sensing *in vivo*.
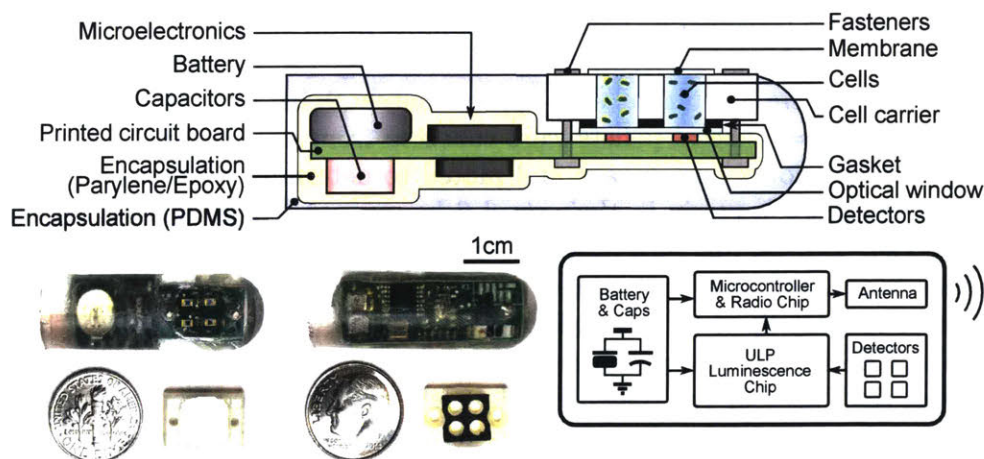
Figure 5-6: **Design of IMBED for miniaturized wireless sensing with cellular biosensors:** Cross section, electrical system diagram, and front and back-side photos of the device.

The electronic system is highly sensitive and captured photon flux down to 5x 10^4 photons/s incident on the 0.29 mm^2 area of the detectors (white-noise limited coefficient of variation 13%$_{rms}$, Fig. 5-7A-B). The mean channel mismatch was less than 6%$_{rms}$ (Fig. 5-7B) and mean temperature-induced drift across 5°C variation was less than 2 pA (Fig. 5-7C). In addition, the electronic system was stable in

simulated gastric fluid for up to 36 h (fig. 5-7D), providing sufficient time to perform an ingestible measurement during gastrointestinal transit.
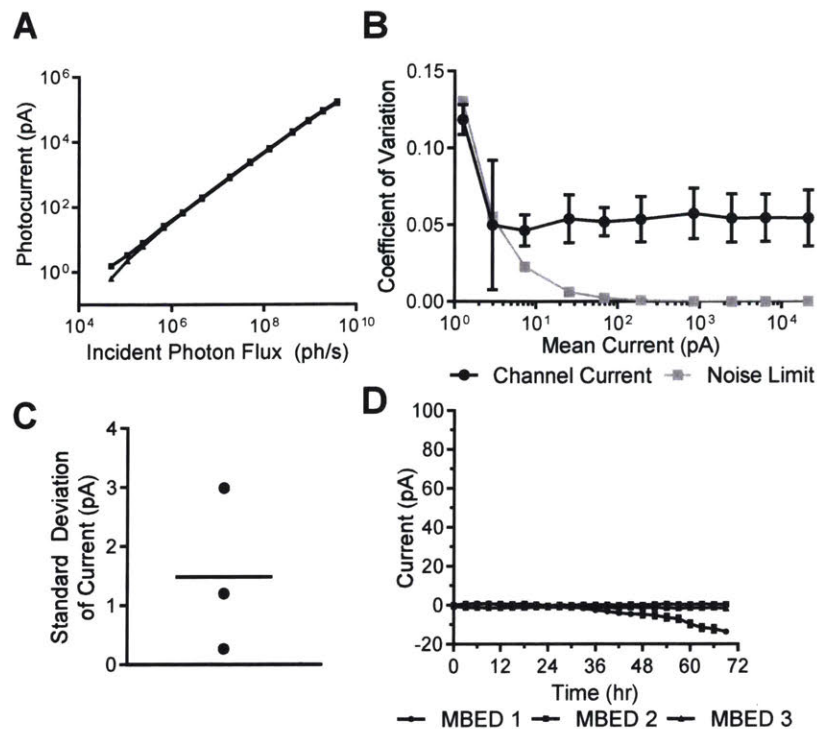
Figure 5-7: **Capsule readout variation was characterized across optical input power, temperature change and fluid submersion:** (A) System photocurrent response measured without cells. The incident photon flux was supplied by green LED (λ = 525nm) and calibrated with an optical power meter (individual traces shown for N = 3 devices). (B) The coefficient of variation between measurements on three channels within a single device, characterized across input light intensity (N = 3 devices). At low signal levels, the measurement standard deviation is limited by white noise ($13\%_{rms}$ noise at 1.3pA). At higher signal levels, it is limited by mismatch between the channels ($< 6\%_{rms}$ above 3pA). (C) Residual variation induced by temperature change, post-calibration. The temperature was stepped from 35°C to 40°C (temperature change 5°C) and the standard deviation across three sensor channels was measured (N = 3 devices). (D) Stability of the measurements from IMBED devices in Simulated Gastric Fluid (SGF) for 72 hours (N = 3). For two devices, current values were stable for the duration of measurement. The third system operated for 36 hours before corruption by humidity became evident.
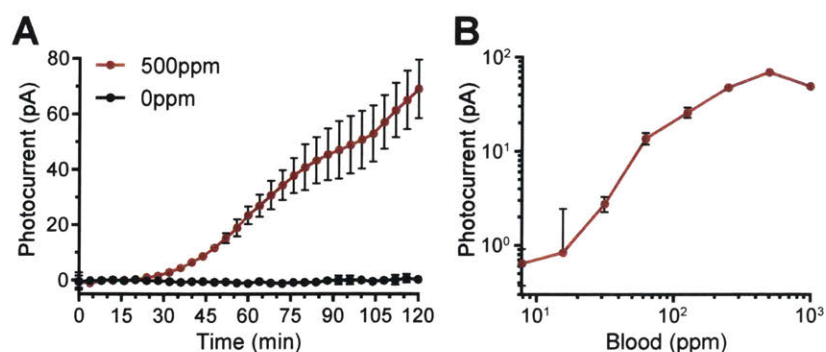
Figure 5-8: *In vitro* **evaluation of IMBEDs with blood biosensors:** (A) Kinetic response of blood sensor IMBED in bacterial growth media supplemented with 0 ppm and 500 ppm blood. (B) Dose-response of blood sensor IMBEDs in bacterial growth media containing different blood concentrations 2 hours post-exposure. The leftmost data point represents the background response in the absence of blood. Error bars denote the SEM for 3 independent biological replicates conducted with different IMBEDs.

5.3.2   *In vitro* Validation of IMBED Functionality

To demonstrate integration of the ingestible luminometer capsule and engineered biosensors, we tested the probiotic blood sensor strains in an IMBED *in vitro*. Upon exposure to 500 ppm blood, induced bioluminescence could be observed as soon as 30 minutes (Fig 2C). This slight delay as compared to plate-reader measurements (fig. S5) likely owes to diffusion time of heme into the cell cavities. The dose-response curve of blood sensor IMBEDs was similar to plate-reader measurements (SNR=76; K_D=135 ppm; compare Fig. 2D and fig. S9), with saturation achieved at 250 ppm and significant detection down to 32.5 ppm blood (Student's t-test; p=0.03). Together, IMBEDs serve as a flexible platform for sensitive detection of bleeding in fluidic environments.

The sensing functionality of IMBEDs can be readily adapted to alternative biomarkers. To illustrate this, we modified thiosulfate and acyl-homoserine lactone (AHL) sensors in bacteria to act as bioluminescent reporters. Thiosulfate could serve as a biomarker of gut inflammation as it is elevated in murine models of colitis [110]. AHLs are molecular signatures of particular bacteria used to coordinate gene expression across populations and their detection in the context of the gut microbiota can

indicate the presence of commensal or infectious agents in the gut [92, 329, 330]. Thiosulfate- and AHL-inducible genetic circuits were introduced into *E. coli* Nissle and exposure to increasing concentrations of inducer led to increasing levels of bioluminescence (Fig. 5-9A-B). When integrated with IMBEDs, biosensing of different analytes was readily detectable in a fluidic environment (Fig. 5-9C). As synthetic biologists continue to develop additional biosensors of clinically-relevant gut biomarkers, we anticipate that the breadth of potential analytes of the IMBED platform will continue to grow.
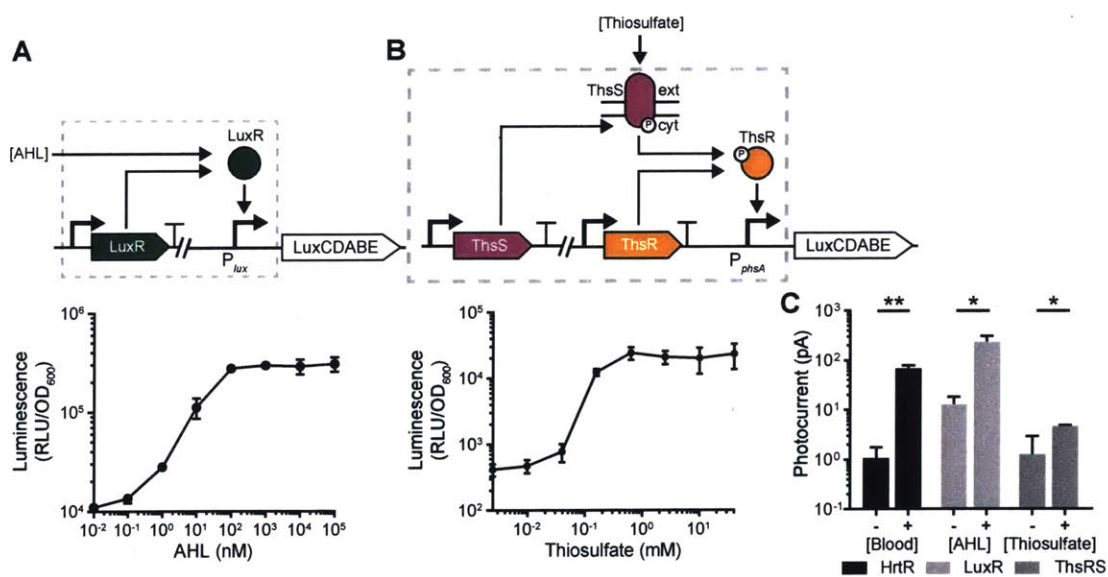
Figure 5-9: **Design and characterization of acyl-homoserine lactone (AHL) and thiosulfate-responsive biosensors:** (A) AHL binds to the transcriptional activator LuxR that activates transcription of the *luxCDABE* operon downstream of the P_{lux} promoter. Titrating increasing amounts of AHL yields high levels of luminescence. (B) The ThsRS two-component system mediated thiosulfate-inducible expression of the *luxCDABE* operon from the P_{phsA} promoter. Thiosulfate binds to the membrane bound ThsS histidine kinase that, in turn, phosphorylates the ThsR response regulator such that it can activate transcription from P_{phsA}. (C) Detection of multiple gut-relevant small molecules with IMBEDs. HrtR-, LuxR- and ThsRS-containing E. coli Nissle strains in IMBEDs were exposed to 500 ppm blood, 100 nM acyl-homoserine lactone (AHL) or 10mM thiosulfate for 2 hours. Error bars indicate SEM from three independent biological replicates. *$P<0.05$, **$P<0.01$, Student's t-test.

5.4 IMBED Performance in Pigs
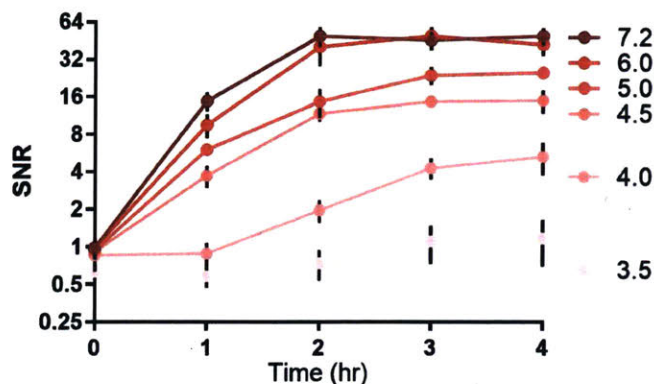
Figure 5-10: **Kinetic response of blood biosensor strains at various pHs:** *E. coli* Nissle blood biosensors (Nissle V2 from Fig. 3-2B) were treated with 500 ppm blood or PBS in citrate-buffered LB at various pHs. Luminescence response was measured in a plate reader every hour for 4 hours. The signal-to-noise ratio (SNR) of each sample at each timepoint was calculated by dividing the OD_{600}-normalized luminescence values of induced by the OD_{600}-normalized luminescence values of uninduced samples. The integrity of the genetic circuit degrades with decreasing pH and is non-functional below pH4.0. Error bars represent SEM of three independent biological experiments.

To examine wireless *in situ* detection of biomolecules with biosensors, we deployed blood sensor IMBEDs in a porcine model of gastrointestinal bleeding. Prior to device deposition, pigs were administered a bicarbonate-glucose neutralization solution with or without 0.25mL of blood (Fig. 5-11A). The neutralization solution helped buffer the low pH of the porcine gastric fluid as acidic environments degrade the functionality of the biosensor (Fig. 5-10). The blood sensor IMBED was subsequently deposited into the stomach via orogastric tube and remained resident and stable in the gastric cavity for the entire duration of the experiment (Fig. 5-11 B and 5-11C). Photocurrent data was wirelessly transmitted from the stomach over the course of 2 hours and logged by both a laptop computer and an Android phone equipped with a custom application for real-time data processing and visualization. The presence of blood in the porcine gastric environment could be observed as early as 52 minutes (Student's t-test; $p<0.05$) and led to a 5-fold increase in photocurrent after 120 minutes as compared to animals given buffer alone (Fig. 5-11D). Luminescence production

was not detected in biosensors lacking the ChuA heme transporter or the luciferase operon, indicating that observed light production was dependent on a functional genetic circuit activated in the presence of heme (Fig. 5-11E). The receiver operating characteristic of the blood sensing IMBED improved over time, with a sensitivity and specificity of 83.3% at 60 minutes and 100% at 120 minutes (Fig. 5-11F). IMBEDs can thus detect low-levels of analyte in the harsh gastric environmental with high specificity and sensitivity.
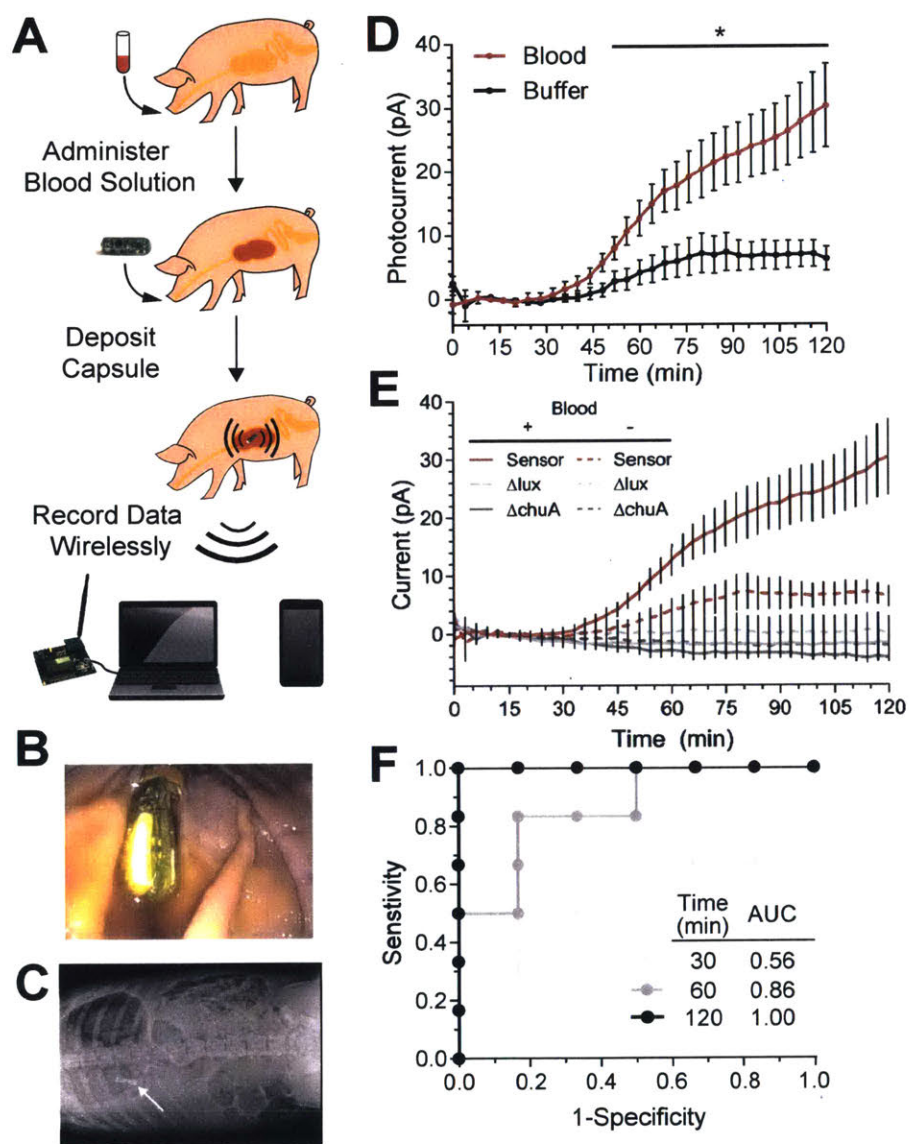
Figure 5-11: **IMBEDs can rapidly detect porcine gastric bleeding.** (A) Schematic of experiment flow which consisted of blood administration in neutralization solution, capsule deposition, and wireless transmission to a commercial receiver connected to a laptop or a cellular phone. Devices were deposited in the stomach of animals administered neutralization solution spiked with blood or with buffer alone. IMBED readings were wirelessly collected for 120 minutes post-device deposition. Representative endoscopic (B) and X-ray (C) images illustrate the location of the device in the stomach at the conclusion of our 2 hour experiments, just prior to device removal (scale bar for C = 5 cm). (D) Kinetic response of blood sensor IMBED in porcine model of gastric bleeding. IMBEDs deposited in gastric cavity can rapidly discriminate between pigs administered blood versus buffer control. *P<0.05, Student's t-test. (E) *E. coli* Nissle strains containing a functional biosensor circuit (Sensor), a circuit lacking the luciferase output (Δlux), and a circuit lacking the heme transporter ChuA (ΔchuA) were loaded into a IMBED. Error bars denote SEM for six IMBED experiments (3 animals on different days, 2 capsules per animal). (F) Receiver operating characteristic (ROC) of IMBED sensing over time. Perfect detection is achieved at t = 120 minutes.

5.5 Discussion

By combining the environmental resilience and natural sensing properties of bacterial cells with the complex data processing and wireless transmission afforded by ultra-low power microelectronics, we developed a device capable of *in vivo* biosensing in harsh, difficult-to-access environments. Using gastrointestinal bleeding as a proof-of-concept model system, we demonstrate strategies for genetic circuit design and optimization, fabrication of an ingestible low-power, wireless luminometer, and validation of integrated system functionality both *in vitro* and in a large animal model. As the field of whole-cell biosensors matures, newly developed sensors of clinically-relevant biomarkers could be rapidly integrated into an IMBED to perform minimally-invasive detection in the gastrointestinal tract. With a test panel of candidate biomolecules, IMBEDs could enable studies of the biochemistry of anatomical regions that are traditionally difficult to access and could lead to the discovery of novel clinical biomarkers associated with health or disease. The *in situ* detection afforded by IMBEDs could also allow sensing of labile gut or microbiota-derived biomolecules that would otherwise be degraded before excretion in stool. Further integration of electronic modules, such as photodetectors, microprocessor, and transmitter, in a

single integrated circuit could allow for further miniaturization of IMBEDs as well as lower power consumption. Additional measurement channels would also enable more precise biochemical readings, as the response of replicate biosensors within the same device could be averaged to mitigate the inherent variance of biological sensors as well as the heterogeneity of the complex gastrointestinal environment. Improved preparation of bacterial cultures for long-term storage, such as lyophilisation, could be implemented to extend the shelf-life of fully assembled IMBEDs. Furthermore, the devices could be equipped with new orally-delivered encapsulation technologies to enable long-term residency, monitoring and anatomic localization in the gastrointestinal tract [331, 332]. This integration of biological engineering and semiconductor electronics offers opportunities to transform diagnosis, management, and monitoring of health and disease.

5.6 Experimental Details

Bacterial Strains and Culture Conditions

Routine cloning and plasmid propagation was performed in *E. coli* DH5α. Gene circuits were initially prototyped in *E. coli* MG1655 and were transferred into probiotic *E. coli* Nissle 1917 for capsule and *in vivo* experiments. Cells were routinely cultured at 37°C in Luria-Bertani (LB) media (Difco). Where appropriate, growth media was supplemented with antibiotics at the following concentrations: 30 μg/mL kanamycin, 100 μg/mL carbenicillin, 25 μg/mL chloramphenicol and 100 μg/mL spectinomycin.

Genetic Part and Plasmid Construction

Genetics parts and plasmids used in this study are listed in Table S2 and Table S3 and will be available from Addgene upon publication. All plasmids were constructed by combining PCR fragments generated by Kapa HiFi Polymerase using Gibson Assembly [188]. Assembly products were transformed into chemically competent *E. coli*

DH5α [191] and sequences were confirmed using Sanger sequencing. Ribosome binding sites (RBSs) of variable strengths were computationally designed using the Salis lab RBS calculator [241, 324].

Growth and Induction

For genetic circuit characterization, overnight cultures were diluted 1:100 in fresh LB and incubated with shaking at 37°C for 2 hours. Cultures were removed from the incubator and 200 μL of culture was transferred to a 96-well plate containing various concentrations of inducer. The plate was returned to a shaking incubator at 37°C. Following 2 hours of incubation, luminescence was read using a BioTek Synergy H1 Hybrid Reader using a 1s integration time and a sensitivity of 135. Luminescence values, measured in relative luminescence units (RLUs), were normalized by the optical density of the culture measured at 600 nm. For *in vitro* kinetic studies, subcultured cells were mixed with inducer in a 96-well plate and immediately placed in the plate reader set at 37°C without shaking. Luminescence and absorbance was read at 5 minute intervals. For pH sensitivity studies, citrate-buffered LB at various pHs was made by supplementing LB media with appropriate volumes of 0.2M sodium phosphate dibasic (Sigma) and 0.1M citric acid (Sigma) and adjusted to the desired pH with HCl. A stock solution of hemin (Sigma) was prepared by dissolving hemin powder in 1M NaOH (Sigma) to a concentration of 25 mM, diluting with double distilled water to a final concentration of 500 μM and sterilizing with a 0.2 μm polyethersulfone (PES) filter. Defibrinated horse blood (Hemostat) was used as the source of blood for most experiments. Blood was lysed by first diluting 1:10 in simulated gastric fluid (SGF) (0.2% NaCl, 0.32% pepsin, 84 mM HCl, pH 1.2) before further dilution in culture media. Stock solutions of sodium thiosulfate (Sigma) and 3-O-C6-HSL (referred to as acyl homoserine lactone (AHL)) (Cayman Chemical) were made in double distilled water.

Indomethacin Mouse Experiments

All mouse experiments were approved by the Committee on Animal Care at the Massachusetts Institute of Technology. Specific-pathogen free (SPF), male C57BL/6J mice (8-10 weeks of age) were purchased from Jackson Labs and were housed and handled under conventional conditions. Mice were acclimated to the animal facility 1 week prior to the commencement of experiments. Animals were randomly allocated to experimental groups. Researchers were not blinded to group assignments. Prior to indomethacin experiments, a pilot experiment was conducted to determine the transit rate of bacteria through the mouse gastrointestinal tract (Figure S6A). Overnight cultures of *E. coli* Nissle were centrifuged at 5000g for 5 minutes and resuspended in an equal volume of 20% sucrose. Animals were inoculated with 200 μL of bacteria culture (approximately 2×10^8 CFU) by oral gavage. Fecal pellets were collected 2, 4, 6, 8, and 24 hours post-gavage, weighed, and homogenized in 1mL of PBS with a 5 mm stainless steel bead using a TissueLyser II (Qiagen) at 25 Hz for 2 minutes. Samples were centrifuged at 500xg for 30 seconds to pellet large fecal debris. Supernatant was serially diluted in sterile PBS and spot plated on MacConkey agar supplemented with kanamycin. Colonies were enumerated following overnight incubation at 37°C. For luminescence assays, luminescence in fecal homogenate was measured in a Biotek Synergy H1 Hybrid Reader with an integration time of 1 second and a sensitivity of 150. Luminescence values were normalized to stool weight normalized CFU values and reported in RLU/CFU. For indomethacin experiments, animals were inoculated with blood sensor bacteria and fecal pellets were collected 6 hours later for luminescence analysis and CFU enumeration. Indomethacin (Sigma) solution was prepared by dissolving the compound in absolute ethanol to a concentration of 20 mg/mL. Immediately prior to mouse gavage, the indomethacin stock solution was diluted to 1.25 mg/mL in PBS and 0.2 mL of dilute indomethacin solution was administered to each animal (10 mg/kg). Preparation of indomethacin solution using this method was essential to ensure reliable and reproducible induction of gastrointestinal bleeding. The following morning, gastrointestinal bleeding was confirmed by performing a

guaiac test (Hemoccult, Beckman Coulter) on fecal pellets from each animal. All mice administered indomethacin were guaiac positive, whereas those administered a PBS control were uniformly guaiac negative. Subsequently, mice were again administered blood sensor bacteria and fecal pellets were collected 6 hours later for luminescence analysis and CFU enumeration.

Preparation of capsules

The electronics in the capsules consisted of four phototransistor detectors (SFH3710, Osram Opto Semiconductors GmbH), a custom bioluminescence detector chip fabricated in a TSMC 65 nm process [313], a microcontroller and radio chip (PIC12LF1840T39A, Microchip Technology Inc.), 22 MHz crystal resonator (7M-22.000MEEQ-T, TXC Corporation), 915 MHz chip antenna (0915AT43A0026, Johanson Technology Inc.), two 220 μF ceramic capacitors (CL32A227MQVNNNE, Samsung Electro-Mechanics America, Inc.), and a 5 mAh lithium manganese button-cell battery (MS621FE-FL11E, Seiko Instruments Inc.). The electronics were soldered onto custom four-layer printed circuit boards (Advanced Circuits Inc.) and two screws were epoxied into mounting holes for later attachment of the plastic cell carriers. The assembly was coated with 4-15 μm of Parylene C to act as a moisture barrier (additional methods describing Parylene C deposition described below). A clear rectangular polycarbonate window (500 μm thickness, Rowland Technologies Inc.) was epoxied above the four phototransistor detectors to provide a flat optical interface. The boards were coated with 1-3 mm of epoxy (20845, Devcon) for mechanical stability and then casted into PDMS capsules 13mm in diameter (Sylgard 184, Dow Corning).

Parylene C deposition

Electronic components were coated in Parylene-C to provide necessary humidity resilience for the sensitive picoampere-level photocurrent measurements. Di-chloro-di-p-xylylene (brand name: diX C) dimer was purchased from Daisan Kasei Co. (now

a KISCO partner company). Thin film Parylene C coating was preformed using an in-house pyrolysis CVD coating tool. After loading the capsules, 10 grams of dimer was loaded into a thermal evaporation heater and the system was evacuated to 1.3 μbar. The pyrolysis furnace and all other vacuum components were pre-heated prior to deposition. During deposition the dimer was evaporated between 105°C to 120°C in order to maintain a constant deposition rate of around 3 $\mathring{A}$/s. Upon reaching the desired thickness the deposition chamber was isolated, the system was cooled, the deposition chamber was vented, and the capsules were removed.

Preparation of cell carriers

Cell carriers were machined or injection-molded in ABS plastic (Protolabs Inc.). Semipermeable membranes (0.22 μm pore size, EIMF22205, Millipore Sigma) were affixed to one side of the cell carriers via heat sealing for 35-45 seconds at 230°C with a stainless steel die. Rubber gaskets for fluidic sealing were die-cut from 380 μm silicone rubber (86435K13, McMaster-Carr) and epoxied to the opposite side of the cell carriers to provide a seal between the carrier and the optical window during experiments.

System operation, packet transmission and reception

The NPN phototransistor detectors were operated in a charge-integration mode using each device's intrinsic capacitance as the charge storage mechanism (measured capacitance, $C_o = 8.7$ nF). The collector of each detector was connected to the supply rail of the system and the emitters were connected to the system ground through independent low-leakage switches (one per detector) in the custom integrated circuit. At the beginning of a measurement, the emitters were shorted to the system ground via the switches and device capacitances were charged to the system voltage. Then, switches were opened and emitter voltages would start to increase independently in response to the dark currents and photo currents in each detector. The custom integrated circuit contained a low-power voltage reference ($V_R = 0.625$ V) and local

oscillator counter (oscillator period, TOSC = 5 ms). In each oscillator cycle, the detector voltages for each channel were compared to the reference voltage and, if the reference was exceeded, a count value was saved corresponding to the number of oscillator cycles required the charge the channel. The on-board microprocessor polled the custom circuit once every 8 seconds to determine whether all four channels had exceeded the reference voltage. Once all were exceeded, the microprocessor read the four counter values through a serial peripheral interface and transmitted a short wireless packet at +10 dBm with count data using an on-board transmitter. Two 220 μF ceramic capacitors included in the capsule supplied the instantaneous peak energy required by the radio transmitter. The data were received externally by a 900 MHz wireless radio (CC1120 Evaluation Kit, Texas Instruments Inc.) attached to a laptop and processed offline in Matlab (The Mathworks, Inc.).

Photocurrent estimation with temperature and offset calibration

In each IMBED, one channel acts as a reference to calibrate for background light and temperature-induced dark current variation, while the remaining three are used for independent measurements. The photocurrent detected by the system was estimated from the measured capsule data using an algorithm for temperature drift and offset calibration, which is described as follows:

Let there be three potentially luminescing sensor channels with counts denoted by $N_i : i = 1, 2, 3$. The time required for the photocurrent stimulated by luminescing cells ($I_{PH,i}$) and the dark background current intrinsic to the photodetectors ($I_{D,i}$) to charge the channel capacitance (C_o) of a channel (i) to the threshold voltage (V_R) was quantized using the number of cycles (N_i) counted by the internal oscillator (period, T_{OSC}). The measured cycles were then used to estimate the photocurrent level. The number of cycles required to charge a sensor channel is given by:

$$N_i = \left(\frac{C_o V_R}{T_{OSC}} \right) \left[\frac{1}{I_{D,i} + I_{PH,i}} \right].$$

Let there be one reference channel containing no luminescing cells ($I_{PH} = 0$) with a count denoted by N_r. The number of cycles required to charge the reference is given by:

$$N_r = \left(\frac{C_o V_R}{T_{OSC}}\right)\left[\frac{1}{I_{D,r}}\right].$$

The desired photocurrent signal on a channel ($I_{PH,i}$) is corrupted by the channel's dark current, which we have modelled as:

$$I_{D,r} = I_{D,OS,i} \cdot f(T),$$

by separating a temperature-independent, channel-specific dark current offset ($I_{D,OS,i}$) from a temperature dependent scaling function [$f(T)$].

To calibrate the temperature and offset, the counts from each sensor channel were first compared to the reference channel by calculating a relative signal R_i:

$$R_i = \frac{1/N_i - 1/N_r}{(1/N_r)} = \left(\frac{I_{D,OS,i} \cdot f(T)}{I_{D,OS,r} \cdot f(T)} - 1\right) + \left[\frac{1}{I_{D,OS,r} \cdot f(T)}\right] I_{PH,i}.$$

In the first term of R_i, the temperature dependence is cancelled, leaving only a dependence on the relative offsets between channels. We can denote this term as $R_{i,OS}$. We used early segments of the count data for each experiment, prior to induction of luminescence from the whole-cell biosensors ($I_{PH,i} = 0$) to estimate $R_{i,OS}$ for each channel. For all experiments, the samples between 0.2 and 0.3 hours (12 to 18 minutes) were used to estimate $R_{i,OS}$. By substituting the measured offset ($R_{i,OS}$), as well as the expression for N_r, we obtained the final expression for the estimated photocurrent in terms of known and measured quantities.

$$I_{PH,i} = \left(\frac{C_O V_R}{T_S N_r}\right)[R_i - R_{i,OS}].$$

This calibration procedure was performed using Matlab software (R2017a, The Mathworks, Inc.).

Optical Calibration

A green LED ($\Lambda = 525$nm, WP7083ZGD/G, Kingbright) was first calibrated across four orders of magnitude of input current using an optical power meter located 30 cm away (PM100D and S130C, Thor Labs Inc.). Three capsules were then placed at the same distance as the power meter and measured across the same LED current conditions. The optical power readings were scaled by the ratio of the area of the phototransistor detectors (0.29 mm^2) to the area of the S130C sensor (70.9 mm^2) in order to estimate the optical power incident on the detectors.

Mobile phone "app" for real-time reception and visualization of results

A 900 MHz USB dongle (CC1111 USB Evaluation Module Kit, Texas Instruments, Inc.) was attached to an Android mobile phone (Galaxy SIII, SCH-I535, Samsung Electronics Co. Ltd.) running a custom application created in Android Studio (Google, Inc.). Temperature and offset calibration was performed on the phone after receiving the first 18 minutes of data to enable offset calibration and the photocurrent estimate was displayed to the user. The raw data was simultaneously uploaded to a cloud service for later analysis.

In vitro IMBED Experiments

LB culture media supplemented with or without inducer (500 ppm lysed blood (unless otherwise noted), 10mM thiosulfate, or 100 nM AHL) was pre-warmed for at least 2 hours prior to the start of experiments. For blood sensor experiments, overnight cultures were diluted 1:10 in 2xYTPG (20g tryptone, 5g NaCl, 10g yeast extract, 22 mL of 1 M potassium phosphate monobasic, 40 mL of 1 M potassium phosphate dibasic, 0.2% glucose, pH 7.2) and 15 μL of diluted culture was added to wells in the cell carrier (approximately 10^6 cells per well). Wild-type *E. coli* Nissle 1917 was added in the reference channel for all experiments. Blood sensor bacteria were added in triplicates to three wells in a single device and values from these three

channels were averaged to obtain a single replicate plotted in Figures 2C-E. Technical replicates are depicted in Figure S8. For thiosulfate and AHL experiments, overnight cultures of ThsRS or LuxR containing cells were subcultured for 2 hours in LB prior to addition to cell carriers. Once all four channels were loaded, the cell carrier was fastened to the capsule and fully submerged in pre-warmed media. Cultures were wrapped several times in thick black fabric to block external light, placed in an incubator at 37°C and data was collected wirelessly for 2 hours. At the end of the experiment, devices were dissembled and cell carriers were discarded. Capsules were sterilized with 70% ethanol and thoroughly washed with distilled water. Capsules were left to air-dry and re-used for future experiments.

Pig Experiments

All pig experiments were approved by the Committee on Animal Care at the Massachusetts Institute of Technology. Female Yorkshire pigs (50-95kg) were obtained from Tufts University and housed under conventional conditions. Animals were randomly selected for the experiments. The animals were placed on a clear liquid diet for 24 hours prior to the experiment with the morning feed held on the day of the experiment. At the time of the experiment, the pigs were sedated with Telazol (tiletamine/zolazepam 5 mg/kg), xylazine (2 mg/kg) and atropine (0.04 mg/kg). An endoscopic overtube (US endoscopy) was placed in the esophagus under endoscopic (Pentax) visual guidance during esophageal intubation. Prior to deposition of devices, 250 mL of neutralization solution (1% sodium bicarbonate and 0.2% glucose) with or without 0.25mL of pig blood was administered directly to the stomach through the endoscope. Overnight bacterial cultures were diluted 1:10 in 2xYTPG and 15 μL of diluted culture was added to wells in the cell carriers. Devices were assembled and deposited in the pig gastric cavity via endoscopic overtube. Full submersion in gastric fluid was confirmed by endoscopic observation. For 2 hours, data from deposited capsules was acquired via a 900 mHz radio attached to a laptop or the Android cellular phone. Endoscopic videos and radiographs of capsules inside the pig stomach were acquired. Devices were retrieved from the gastric cavity using a hexagonal snare. A

total of 6 animals were included in the experiments; 3 were administered neutralization solution containing blood and 3 served as negative controls. Two devices were deposited per pig, such that each group has a sample size of 6.

Data Analysis, Statistics and Computational Methods

All data were analyzed using GraphPad Prism version 7.03 (Graph Software, San Diego, CA, USA, http://www.graphpad.com). Sequence analysis was performed using Geneious version 9.1.8 (http://www.geneious.com) [333]. As noted, error bars represent the SEM of at least three independent experiments carried out on different days. Significance between groups was determined using an unpaired, two-tailed Student's t-test assuming unequal variance. Fold change or signal-to-noise ratio was determined by dividing the normalized luminescence values (RLU/CFU) of samples treated with the maximal inducer concentration with uninduced samples. Response curves were fit to a Hill function: $Y=(B_{max}X^n)/(K^n+X^n) + C$, where X is the inducer concentration, Y is the normalized luminescence output, Bmax is the maximum luminescence, K is the threshold constant, n is the Hill coefficient and C is the baseline luminescence.

Chapter 6

Conclusion

6.1 Overview

The profound impact that microbial communities have on human health is providing new diagnostic and therapeutic avenues for treating disease. However, existing therapeutic approaches for modulating microbiomes in the clinic remain relatively crude. Exciting research into additive, subtractive or modulatory strategies to affect the human microbiota and, in turn, human health, are progressing towards the clinic, powered by advancements in synthetic biology and microbial ecology. In this thesis, I contributed to the field of microbiome engineering through the development of novel subtractive and additive approaches to manipulating microbial communities.

Text and concepts in the following coclusion have been adapted from the 2016 Advanced Drug Delivery review entitled 'Microbiome therapeutics – Advances and Challenges' to which I contributed as a first author [2].

6.2 Major Contributions

The various works elaborated in this thesis have made strides in the field of microbiome engineering:

6.2.1 Subtractive Approach: Engineering Bacteriophage for Sequence Specific Antimicrobials

Precision editing of microbial communities could enable reductionist studies to elucidate the role of specific microbes in community composition and in host physiology. However, the majority of current techniques to remove bacteria from communities are broad spectrum in nature and remove large swaths of unrelated bacteria. In Chapter 2, I described an approach of using bacteriophage equipped with CRIPSR-Cas machinery to selectively eliminate strains from a bacterial population based on the presence of target DNA sequences. CRIPSR-Cas was programmed to cause lethal double stranded breaks in antibiotic resistance and virulence genes. The engineered phage particles could discriminate between closely related strains in a model bacterial consortium. The work demonstrates the potential utility of engineered bacteriophage to act as selective antimicrobials in the context of host-associated microbial communities.

6.2.2 Additive Approach: Genetic Toolset to Domesticate Commensal *Bacteroides* spp. to Modulate Host Immunity

Engineering commensal microbes for *in situ* sensing and production of therapeutic agents could provide much needed, cost-effective therapies for chronic diseases. Commensals can be engineered to sense biomarkers associated with a disease state and produce therapeutic molecules in response to restore homeostasis. However, engineering commensal organisms is challenging as there are few genetic tools to modify their genomes. In Chapter 3, I described efforts to domesticate commensal *Bacteroides* spp. with a suite of synthetic biology tools that facilitate organism engineering. First, we developed the highly efficient and specific pNBU1 vector system for integration of genetic circuits. Next, we created a library of constitutive and inducible promoters and RBSs for precision gene expression in *B. thetaiotaomicron*. We then

use these genetic parts to build recombinase-based genetic memory switches and re-programmable CRIPSRi-based repression systems. Finally, we validate functionality of these engineered strains in the context of the murine microbiota. The work described in this chapter are a valuable resource for *Bacteroides* engineering and have been implemented in later works by other groups [303, 334]. In addition, the work serves as a blueprint that one may follow to domesticate intractable commensals.

In Chapter 4, the engineering toolset described in Chapter 3 was extended to protein secretion in *Bacteroides* spp. Outer membrane vesicles (OMVs) produced by *Bacteroides* are secreted proteoliposomal complexes that can directly interact with the host immune system. Rational modification of OMVs could therefore be a powerful approach to engineer immune responses in the gut. Using proteomic and bioinformatics approaches, we found that *Bacteroides* OMVs were enriched in membrane-anchored lipoproteins. Translational fusion to both full-length lipoproteins and lipoprotein secretion signal sequences was sufficient to tether heterologous proteins to OMVs derived from several species in the *Bacteroides* genus. Using *Bacteroides* Vesicle Incorporation Tags (BVITs), we packaged mammalian cytokines and model antigens in OMVs with the goal of modulating gut immune responses. Antigen-loaded OMVs could be recognized by antigen presenting cells and induce T cell proliferation and high levels of IL-10 secretion. Antigen-loaded OMVs could thus be an effective tool to manipulate antigen-specific immune responses in autoimmune or allergic diseases. The development and characterization of a prominent secretion mechanism in *Bacteroides* could also serve as a resource for heterologous protein secretion from these commensal organisms.

6.2.3 Additive Approach: Engineered Biosensors to Measure Gastrointestinal Biomarkers *in situ*

Commensal organisms naturally respond to changes in environmental changes with alterations in gene expression. Capturing this dynamic response to perturbations could be an effective means to detect biomarkers associated with disease states.

However, real-world applications of cellular biosensors are hampered by difficulties in gene expression analysis, which requires biological sampling, processing and expensive laboratory techniques. In Chapter 4, I described collaborative efforts between the domains of biological and electrical engineering that yielded a hybrid bio-electronic medical device for *in situ* detection of gastrointestinal biomarkers. As a proof-of-concept, we developed a probiotic *E. coli* heme biosensor that could sense bleeding in the murine gastrointestinal tract. When combined with a custom-designed miniaturized luminometer, we demonstrated *in vitro* detection of blood in fluidic samples as well as detection in the porcine gastric cavity. Integration of biosensors for alternative biomarkers could be readily integrated into the device, demonstrating broad utility for *in situ* biomarker detection. This work progresses the use of cellular biosensors towards clinical management of gastrointestinal diseases.

6.3   Future Directions

The works described herein reflect contributions that push microbiome engineering forward and hopefully capture the potential and excitement surrounding the field. However, it should be evident that the goal of precision editing of the microbiota has yet to be achieved. Further work in multiple areas of research are necessary to push the promise of microbiota engineering to clinical applications. I have previously defined and elaborated on these major challenges in greater detail in [2]. Fundamental understanding of the forces that shape host-associated microbial communities and mediate host-bacterial interactions is essential for the rational design of microbiome therapeutics. The development of clinically relevant biosensors and stable, robust genetic circuits would help realize the vision of fully autonomous cellular therapies. Finally, cognisance of regulatory, biocontainment and safety issues during research of novel therapeutics should help hasten translation of basic research into the clinic.

6.3.1 Subtractive Challenges: Engineering Bacteriophage Host-Range

While the CRISPR-based antimicrobial approach in Chapter 2 could serve as a powerful tool to edit bacterial communities, it is currently restricted by the limited host range of bacteriophage. In natural settings, phage host range is generally restricted to a subset of strains within a single species, although there are some descriptions of phages that can cross the species, genus, and even family barriers [335] . Synthetic biology, however, has rendered the manipulation of bacteriophage genomes more accessible [138, 336] and biological engineering strategies may be able to overcome host range limitations. Rational or random modifications of host range determinants on a malleable phage scaffold could broaden or shift host range [138]. Using broad host range bacteriophages as scaffolds for engineering may also prove to yield a more versatile platform for gene delivery. Incorporation of components that inhibit bacterial phage defense systems, such as anti-CRISPRs [337] and anti-restriction strategies [338], may also be employed to improve the efficacy of phage in a wider range of bacteria.

Despite the promise of bacteriophage as therapeutics, there are currently few examples of successful therapy in large scale clinical trials. Our unpublished personal experience with bacteriophage in mouse models of gut decolonization also suggests that *in vitro* phage performance poorly predicts success *in vivo*. There are currently few studies that detail the differences between phage replication *in vivo* vs *in vitro*. Differences in biogeography of phages and their prey, the growth state of infected bacteria, the chemical composition of body sites, infection on non-clonal bacterial populations, bacterial host evolution, and phage bioavailability may affect the viability of phage therapy. Some studies have suggested an integral role of the host immune system in aiding bacteriophage in pathogen clearance in lung models of infection [339], implying *in vivo* phage infection should be viewed through an interkingdom lens. Characterization of *in vivo* phage replication will no doubt be instrumental in informing engineering strategies to improve therapeutic phage preparations.

6.3.2 Additive Challenges: Stable Engraftment of Engineered Organisms

A comprehensive understanding of the rules which govern invasion, resilience and succession in a host-associated microbial ecosystem is crucial for the development of long-term cell-based therapies. Indeed, for some applications, stable colonization of recombinant microbes or microbial communities may not be necessary for treatment of disease if cells can exert their intended function during transit through the intestine. For instance, many current efforts for recombinant therapies have employed *L. lactis*, a bacterium which does not colonize the mammalian intestine, as a chassis for therapeutic protein production [95–97, 99, 103]. However, stable engraftment of cell-based therapies into the endogenous microbiota may increase the efficacy of treatment and enable the development of long-term, fully autonomous, disease-responsive therapies.

Further studies into the ecological principles that underpin the human microbiome are necessary for rational manipulation of microbial communities. Importantly, pairing additive approaches with subtractive or modulatory ones could yield more reliable and stable engraftment into the microbiome. Targeted antimicrobials, such as bacteriophage, or dietary supplementation with prebiotics could open niches for therapeutic microbes to occupy. As the gut microbiota is a complex ecological community, assimilating these biological findings into predictive systems biology models [340, 341] will greatly facilitate the creation of microbiota-based therapeutics.

Different therapeutic applications should demand different chassis organisms for cell-based therapies. To date, *E. coli* Nissle 1917 [91, 104], *L. lactis* [95–97, 99, 103], *Lactobacillus* spp. [93, 101, 105] and *Bacteroides* spp. [100, 196] have served as vectors for cell-based therapies. While *L. lactis* cannot colonize the intestine, *E. coli* and *Lactobacillus* spp. tend to be enriched in the small intestine, whereas *Bacteroides* spp. reside in the cecum and colon [342]. Some species tend to colonize the mucosal layer, whereas others prefer the intestinal lumen [343–345]. Some strains may also colonize some hosts better than others [109]. The biogeography of disease should

thus dictate the best suited organism for therapy. The natural health benefits of bacteria can also serve as an adjuvant for engineered therapeutics. For example, *B. fragilis*, *Faecalibacterium prausnitzii* and those from *Clostridium* clusters IV and XIVa naturally protect against inflammatory disease [58, 61, 346] and may serve as prime candidates for anti-inflammatory therapies compared to *E. coli*, which are enriched in an inflamed gut [347]. Finally, organisms that are naturally abundant and resilient to environmental perturbations should be preferred for long-term therapies. To this end, the choice of microbial chassis should be paired with thorough characterization of its relative benefits for the target disease. The development of methods for engineering currently intractable organisms would also enable additional possibilities for cell-based therapies. While ample genetic tools exist for *E. coli* and lactic acid bacteria and our recent work has extended these to *Bacteroides* spp. [196], new and efficient genetic techniques for the manipulation of intractable group IV and XIVa Clostridia and *F. prausnitzii* are needed to accelerate the development of future therapies that utilize these organisms.

The robustness of engineered functions to time and changing environments is a major challenge to long-term cellular therapies. Currently, prototyping the complex genetic circuits required for sense-and-respond cellular therapeutics most often occurs in optimal in vitro growth conditions that may not accurately replicate the intended environment. As work in synthetic biology progresses towards animal models and clinical applications, new, more sophisticated in vitro systems should be developed to better emulate the variable conditions faced by the endogenous microbiota. Both single [348, 349] and multi-stage [350] chemostats have been used to support culture of fecal samples and could provide a testbed to understand how interbacterial interactions impact functionality of genetic circuits. Additionally, organoid [351], 3-D intestinal scaffolds [352] and gut-on-a-chip [353] models have also been employed to predict interactions between probiotics and the host.

Genetic circuits described in synthetic biology often fail to report the evolutionary stability of engineered functions. Assessment of gene circuit function is generally made on short timescales (<24 hours). However, cellular therapies may need to

retain function for weeks or months to achieve significant efficacy. The numerous genetic parts required to maintain a complex gene circuit are energetically taxing on cells [354]. *In vitro* evolution experiments revealed this strong fitness cost can lead to a rapid loss of function in engineered bacteriophages [355, 356]. Efforts to quantify and minimize the burden placed on recombinant cellular therapies [354] should help maintain long-term function of therapeutics in the competitive context of the microbiota.

6.3.3 Additive Challenges: Development of Clinically Relevant Sensors

The creation of fully autonomous cell-based therapies demands a well-characterized library of biosensors that would allow dynamic responses to environmental perturbations. Synthetic biology has engendered a plethora of genetic parts for sensing and integrating environmental signals into changes in gene expression. Luminescent, fluorescent or colorimetric readouts are common outputs of biosensors and can be designed to be transient if transcriptionally regulated or permanent if coupled to genomic alterations [196, 216, 218, 220] or a bistable toggle switch [109, 357]. Some of these genetic circuits have been implemented *in vivo* for sensing a small-molecule inducer [109], dietary carbohydrates [196, 304], markers of inflammation [110, 111], or cancer [358]. The blood sensor developed in this thesis add to that body of work. To date, the development of biosensors has relied on genome mining and previous descriptions of systems in scientific literature. Our efforts in Chapter 5 adopted this *ad hoc* approach. Directed evolution has been successfully applied to change substrate specificity of enzymes [359] and promoter specificity of RNA polymerases [360] and similar strategies could be applied for the development of novel biosensors. Moreover, hybrid transcription factors composed of distinct DNA-binding and ligand-binding domains has expanded the collection of available sensors [361–363]. A generalized approach for *de novo* discovery of sensors of clinically relevant molecules is greatly needed, as well as fluid dialogue between synthetic biologist and clinicians

to identify relevant candidate disease biomarkers. By pairing metabolomic studies with streamlined strategies for biosensor discovery, engineered microbes could offer a new class of diagnostics by assaying concentrations of biomarkers inside the body instead of in *ex vivo* samples. These sensors could be used to trigger the expression of therapeutic molecules to enable on-demand and localized production of medicines only during active disease.

6.3.4 Regulation, Safety and Biocontainment

As cell-based therapies constitute a novel paradigm in drug development, a regulatory framework that addresses safety and biocontainment issues should be established to minimize adverse events and environmental release of engineered organisms. Many of the bacterial chassis used for recombinant therapies are generally recognized as safe (GRAS) organisms as listed by the U.S. Food and Drug Administration. These organisms are largely probiotics or bacteria employed in the production of food products. However, as natural commensals, such as Clostridial or *Bacteroides* species, may be prime candidates for microbiota-based therapies, the safety of these organisms should be evaluated. The ability of these organisms to stably colonize their target environments may enable greater therapeutic efficacy, but may raise questions about the pharmacology and control of such therapies. The spread of genetically modified DNA from recombinant organisms to endogenous members of the microbiota may also be a concern as natural horizontal gene transfer is prevalent in the human microbiome [364]. The escape of engineered organisms into the environment that may lead to unintentional colonization of others may be of similar concern, even though most genetically modified organisms developed in the lab seem to be less fit than wild-type [354]. Conditional kill switches that eliminate engineered microbes [361, 365] or destroy genetic circuits [366] have been created for biocontainment. Generating auxotrophic microbes that cannot replicate in or outside of the gut has also proved a successful strategy in limiting the spread of recombinant cells [202]. Indeed, auxotrophy was employed as the sole biocontainment strategy in early clinical trials involving recombinant microbes [95, 103]. Further, physical containment of

engineered microbes in polymer shells could provide an extra level of safety to minimize unwanted escape of engineered organisms. Moving forward, a dialogue between regulators and researchers should help shape the creation of technologies necessary for the safe implementation of cellular therapies.

Appendix A

B. *thetaiotaomicron* OMV Proteins Identified by Mass Spectrometry

Gene Designation	Found in Elhenawy 2014	Av Tot Pep	Av UniPep	Secretion Tags	Gene Ontology Description
BT0177	X	25	3	SpI	NigD-like protein
BT1486	X	128	15	N	DUF5074 domain-containing protein
BT1487	X	189	22	SpII	DUF5074 domain-containing protein
BT1488	X	54	9	N	cell surface protein
BT1491	X	112	15	SpII	DUF4465 domain-containing protein
BT1792		20	14	SpII	cell surface protein
BT1895	X	21	3	SpII	Uncharacterised protein
BT1896	X	15	4	N	cell surface protein
BT1938		15	7	N	putative lipoprotein
BT1954	X	14	7	N	putative surface layer protein
BT1955	X	562	34	SpII	glycoside hydrolase xylanase family
BT1956	X	73	17	SpII	cell surface protein
BT2263	X	24	10	SpII	SusD/RagB family nutrient-binding outer membrane lipoprotein
BT2365	X	27	11	N	hypothetical protein
BT2479		29	13	SpII	peptidase M75
BT3066		53	22	SpII	Uncharacterised protein
BT3067		29	10	SpII	Uncharacterised protein
BT3147	X	187	19	N	GLUG domain-containing protein
BT3148	X	29	10	N	conserved domain protein
BT3222	X	29	8	SpII	DUF4848 domain-containing protein
BT3236	X	34	10	SpII	DUF4302 domain-containing protein
BT3237		20	9	SpII	Uncharacterised protein
BT3238	X	45	17	SpII	RagB/SusD family nutrient uptake outer membrane protein
BT3241	X	41	15	SpII	RagB/SusD family nutrient uptake outer membrane protein
BT3312		15	8	SpII	glucosylceramidase
BT3413	X	26	11	SpI	Uncharacterised protein
BT3561	X	26	10	SpII	Uncharacterised protein
BT3562	X	37	20	SpII	carbohydrate binding domain protein
BT3727		16	4	SpII	Uncharacterised protein
BT3740	X	44	16	SpII	major paralogous domain-containing protein
BT3741		43	15	SpII	Uncharacterised protein
BT3742	X	347	28	SpII	Calycin-like beta-barrel domain-containing protein
BT3743		127	19	SpII	DUF5018 domain-containing protein
BT3745		24	2	SpII	Putative carbohydrate metabolism domain-containing protein
BT3960		28	10	SpII	Calpain family cysteine protease
BT4005	X	14	7	SpII	NigD-like protein
BT4306	X	15	9	SpII	DUF4270 domain-containing protein

Bibliography

1. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **486,** 207–214 (2012).

2. Mimee, M., Citorik, R. J. & Lu, T. K. Microbiome therapeutics - Advances and challenges. *Advanced Drug Delivery Reviews* **105,** 44–54 (2016).

3. Dominguez-Bello, M. G. *et al.* Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proceedings of the National Academy of Sciences* **107,** 11971–11975 (2010).

4. Bergstrom, A. *et al.* Establishment of Intestinal Microbiota during Early Life: a Longitudinal, Explorative Study of a Large Cohort of Danish Infants. *Applied and Environmental Microbiology* **80,** 2889–2900 (2014).

5. Cho, I. *et al.* Antibiotics in early life alter the murine colonic microbiome and adiposity. *Nature* **488,** 621–626 (2012).

6. Faith, J. J. *et al.* The long-term stability of the human gut microbiota. *Science (New York, N.Y.)* **341,** 1237439 (2013).

7. Franzosa, E. A. *et al.* Identifying personal microbiomes using metagenomic codes. *Proceedings of the National Academy of Sciences* **112,** E2930–E2938 (2015).

8. Maurice, C. F., Haiser, H. J. & Turnbaugh, P. J. Xenobiotics shape the physiology and gene expression of the active human gut microbiome. *Cell* **152,** 39–50 (Jan. 2013).

159

9. Morgan, X. C. *et al.* Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biology* **13,** R79 (2012).

10. Lozupone, C. a., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489,** 220–230 (2012).

11. Voreades, N., Kozil, A. & Weir, T. L. Diet and the development of the human intestinal microbiome. *Frontiers in Microbiology* **5,** 1–9 (2014).

12. David, L. a. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505,** 559–63 (2014).

13. Carmody, R. N. *et al.* Diet Dominates Host Genotype in Shaping the Murine Gut Microbiota. *Cell Host & Microbe* **17,** 72–84 (2015).

14. Dey, N. *et al.* Regulators of Gut Motility Revealed by a Gnotobiotic Model of Diet-Microbiome Interactions Related to Article Regulators of Gut Motility Revealed by a Gnotobiotic Model of Diet-Microbiome Interactions Related to Travel. *Cell* **163,** 95–107 (2015).

15. Zhang, C. *et al.* Structural resilience of the gut microbiota in adult mice under high-fat dietary perturbations. *The ISME Journal* **6,** 1848–1857 (2012).

16. Falony, G. *et al.* Population-level analysis of gut microbiome variation. *Science* **352,** 560–564 (Apr. 2016).

17. Rothschild, D. *et al.* Environmental factors dominate over host genetics in shaping human gut microbiota composition. *bioRxiv* **555,** 150540 (2017).

18. Bérdy, J. Bioactive Microbial Metabolites. *The Journal of Antibiotics* **58,** 1–26 (2005).

19. Donia, M. S. *et al.* A Systematic Analysis of Biosynthetic Gene Clusters in the Human Microbiome Reveals a Common Family of Antibiotics. *Cell* **158,** 1402–1414 (2014).

20. Cohen, L. J. *et al.* Commensal bacteria make GPCR ligands that mimic human signalling molecules. *Nature* **549,** 48–53 (2017).

21. Guo, C. J. *et al.* Discovery of Reactive Microbiota-Derived Metabolites that Inhibit Host Proteases. *Cell* **168,** 517–526.e18 (2017).

22. Smanski, M. J. *et al.* Functional optimization of gene clusters by combinatorial design and assembly. *Nature Biotechnology* **32** (2014).

23. Ley, R. E. *et al.* Obesity alters gut microbial ecology. *Proceedings of the National Academy of Sciences of the United States of America* **102,** 11070–11075 (2005).

24. Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444,** 1027–1031 (2006).

25. Ridaura, V. K. *et al.* Gut Microbiota from Twins Discordant for Obesity Modulate Metabolism in Mice. *Science* **341,** 1241214 (Sept. 2013).

26. Smith, M. I. *et al.* Gut Microbiomes of Malawian Twin Pairs Discordant for Kwashiorkor. *Science* **339,** 548–554 (Feb. 2013).

27. Yang, J.-Y. *et al.* Gut commensal Bacteroides acidifaciens prevents obesity and improves insulin sensitivity in mice. *Mucosal Immunology* (2016).

28. Plovier, H. *et al.* A purified membrane protein from Akkermansia muciniphila or the pasteurized bacterium improves metabolism in obese and diabetic mice. *Nature Medicine* **23,** 107–113 (2016).

29. Pedersen, H. K. *et al.* Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature,* 1–6 (2016).

30. Canfora, E. E., Jocken, J. W. & Blaak, E. E. Short-chain fatty acids in control of body weight and insulin sensitivity. *Nature Reviews Endocrinology* **11,** 577–591 (2015).

31. Blouin, J.-M. *et al.* Butyrate elicits a metabolic switch in human colon cancer cells by targeting the pyruvate dehydrogenase complex. *International Journal of Cancer* **128,** 2591–2601 (2011).

32. Donohoe, D. R. *et al.* The Microbiome and Butyrate Regulate Energy Metabolism and Autophagy in the Mammalian Colon. *Cell Metabolism* **13**, 517–526 (2011).

33. Kelly, C. J. *et al.* Crosstalk between Microbiota-Derived Short-Chain Fatty Acids and Intestinal Epithelial HIF Augments Tissue Barrier Function. *Cell Host & Microbe* **17**, 662–671 (2015).

34. Davie, J. R. Inhibition of histone deacetylase activity by butyrate. *The Journal of nutrition* **133**, 2485S–2493S (July 2003).

35. Brown, A. J. *et al.* The Orphan G protein-coupled receptors GPR41 and GPR43 are activated by propionate and other short chain carboxylic acids. *The Journal of biological chemistry* **278**, 11312–9 (2003).

36. Samuel, B. S. *et al.* Effects of the gut microbiota on host adiposity are modulated by the short-chain fatty-acid binding G protein-coupled receptor, Gpr41. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 16767–16772 (2008).

37. De Vadder, F. *et al.* Microbiota-generated metabolites promote metabolic benefits via gut-brain neural circuits. *Cell* **156**, 84–96 (Jan. 2014).

38. Tolhurst, G. *et al.* Short-Chain Fatty Acids Stimulate Glucagon-Like Peptide-1 Secretion via the G-Protein-Coupled Receptor FFAR2. *Diabetes* **61**, 364–371 (2012).

39. Soret, R. *et al.* Short-Chain Fatty Acids Regulate the Enteric Neurons and Control Gastrointestinal Motility in Rats. *Gastroenterology* **138**, 1772–1782 (2010).

40. Koeth, R. A. *et al.* γ-Butyrobetaine Is a Proatherogenic Intermediate in Gut Microbial Metabolism of L-Carnitine to TMAO. *Cell Metabolism* **20**, 799–812 (2014).

41. Wang, Z. *et al.* Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**, 57–63 (2011).

42. Gregory, J. C. *et al.* Transmission of Atherosclerosis Susceptibility with Gut Microbial Transplantation. *Journal of Biological Chemistry* **290**, 5647–5660 (2015).

43. Wang, Z. *et al.* Non-lethal Inhibition of Gut Microbial Trimethylamine Production for the Treatment of Atherosclerosis. *Cell* **163**, 1585–1595 (2015).

44. Maier, L. *et al.* Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature* **555**, 623–628 (2018).

45. Forslund, K. *et al.* Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* **528**, 262–266 (2015).

46. Imhann, F. *et al.* Proton pump inhibitors affect the gut microbiome. *Gut* **65**, 740–748 (May 2016).

47. Jackson, M. A. *et al.* Proton pump inhibitors alter the composition of the gut microbiota. *Gut* **65**, 749–756 (May 2016).

48. Dobkin, J. F., Saha, J. R., Butler, V. P., Neu, H. C. & Lindenbaum, J. Digoxin-inactivating bacteria: identification in human gut flora. *Science (New York, N.Y.)* **220**, 325–7 (Apr. 1983).

49. Haiser, H. J. *et al.* Predicting and Manipulating Cardiac Drug Inactivation by the Human Gut Bacterium Eggerthella lenta. *Science* **341**, 295–298 (July 2013).

50. Wallace, B. D. *et al.* Alleviating cancer drug toxicity by inhibiting a bacterial enzyme. *Science* **330**, 831–5 (2010).

51. Belkaid, Y. & Hand, T. W. Role of the Microbiota in Immunity and Inflammation. *Cell* **157**, 121–141 (2014).

52. Kabat, A. M., Srinivasan, N. & Maloy, K. J. Modulation of immune development and function by intestinal microbiota. *Trends in Immunology* **35**, 507–517 (2014).

53. Round, J. L. & Mazmanian, S. K. The gut microbiota shapes intestinal immune responses during health and disease. *Nature reviews. Immunology* **9**, 313–323 (2009).

54. Wlodarska, M., Kostic, A. D. & Xavier, R. J. An Integrative View of Microbiome-Host Interactions in Inflammatory Bowel Diseases. *Cell Host & Microbe* **17**, 577–591 (2015).

55. Geva-Zatorsky, N. *et al.* Mining the Human Gut Microbiota for Immunomodulatory Organisms. *Cell* **168**, 928–943.e11 (2017).

56. Vatanen, T. *et al.* Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans. *Cell,* 1–12 (Apr. 2016).

57. Haberman, Y. *et al.* Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. *The Journal of clinical investigation* **124**, 3617–33 (2014).

58. Mazmanian, S. K., Round, J. L. & Kasper, D. L. A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* **453**, 620–5 (May 2008).

59. Round, J. L. & Mazmanian, S. K. Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 12204–9 (July 2010).

60. Arpaia, N. *et al.* Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature* **504**, 451–5 (2013).

61. Atarashi, K. *et al.* Induction of Colonic Regulatory T Cells by Indigenous Clostridium Species. *Science* **331**, 337–341 (Jan. 2011).

62. Furusawa, Y. *et al.* Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **504**, 446–50 (2013).

63. Ivanov, I. I. *et al.* Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* **139**, 485–98 (2009).

64. Yang, Y. *et al.* Focused specificity of intestinal TH17 cells towards commensal bacterial antigens. *Nature* **510,** 152–6 (2014).

65. Arrieta, M.-C. *et al.* Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Science Translational Medicine* **7,** 307ra152–307ra152 (Sept. 2015).

66. Stefka, A. T. *et al.* Commensal bacteria protect against food allergen sensitization. *Proceedings of the National Academy of Sciences* **111,** 13145–13150 (2014).

67. Buffie, C. G. & Pamer, E. G. Microbiota-mediated colonization resistance against intestinal pathogens. *Nature Reviews Immunology* **13,** 790–801 (2013).

68. Sassone-Corsi, M. & Raffatellu, M. No Vacancy: How Beneficial Microbes Cooperate with Immunity To Provide Colonization Resistance to Pathogens. *The Journal of Immunology* **194,** 4081–4087 (2015).

69. Rupnik, M., Wilcox, M. H. & Gerding, D. N. Clostridium difficile infection: new developments in epidemiology and pathogenesis. *Nature Reviews Microbiology* **7,** 526–536 (2009).

70. Deriu, E. *et al.* Probiotic Bacteria Reduce Salmonella Typhimurium Intestinal Colonization by Competing for Iron. *Cell Host & Microbe* **14,** 26–37 (July 2013).

71. Buffie, C. G. *et al.* Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. *Nature* **517,** 205–208 (Oct. 2014).

72. Russell, A. B. *et al.* A Type VI Secretion-Related Pathway in Bacteroidetes Mediates Interbacterial Antagonism. *Cell Host & Microbe,* 1–10 (July 2014).

73. Ferreyra, J. A. *et al.* Gut Microbiota-Produced Succinate Promotes C. difficile Infection after Antibiotic Treatment or Motility Disturbance. *Cell Host & Microbe* **16,** 770–777 (2014).

74. Ng, K. M. *et al.* Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. *Nature* **502,** 96–99 (2013).

75. Lyte, M. & Cryan, J. F. *Microbial Endocrinology : The Microbiota- Gut-Brain Axis in Health and Disease* (eds Lyte, M. & Cryan, J. F.) (Springer, 2014).

76. Sampson, T. R. & Mazmanian, S. K. Control of Brain Development, Function, and Behavior by the Microbiome. *Cell Host & Microbe* **17**, 565–576 (2015).

77. Sharon, G. *et al.* Perspective Specialized Metabolites from the Microbiome in Health and Disease. *Cell Metabolism* **20**, 719–730 (2014).

78. Sampson, T. R. *et al.* Gut Microbiota Regulate Motor Deficits and Neuroinflammation in a Model of Parkinson's Disease. *Cell* **167**, 1469–1480.e12 (2016).

79. Yano, J. M. *et al.* Indigenous Bacteria from the Gut Microbiota Regulate Host Serotonin Biosynthesis. *Cell* **161**, 264–276 (2015).

80. Williams, B. B. *et al.* Discovery and Characterization of Gut Microbiota Decarboxylases that Can Produce the Neurotransmitter Tryptamine. *Cell Host & Microbe,* 1–9 (Sept. 2014).

81. Hsiao, E. Y. *et al.* Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell* **155**, 1451–63 (Dec. 2013).

82. Derrien, M. & van Hylckama Vlieg, J. E. Fate, activity, and impact of ingested bacteria within the human gut microbiota. *Trends in Microbiology* **23**, 354–366 (2015).

83. Frei, R., Akdis, M. & O'Mahony, L. Prebiotics, probiotics, synbiotics, and the immune system. *Current Opinion in Gastroenterology* **31**, 153–158 (2015).

84. De LeBlanc, A. d. M. Effect of probiotic administration on the intestinal microbiota, current knowledge and potential applications. *World Journal of Gastroenterology* **20**, 16518 (2014).

85. Marchesi, J. R. *et al.* The gut microbiota and host health: a new clinical frontier. *Gut,* 1–10 (2015).

86. Varankovich, N. V., Nickerson, M. T. & Korber, D. R. Probiotic-based strategies for therapeutic and prophylactic use against multiple gastrointestinal diseases. *Frontiers in Microbiology* **6,** 1–14 (2015).

87. Cuello-Garcia, C. a. *et al.* Probiotics for the prevention of allergy: A systematic review and meta-analysis of randomized controlled trials. *Journal of Allergy and Clinical Immunology,* 1–10 (2015).

88. Fujiya, M., Ueno, N. & Kohgo, Y. Probiotic treatments for induction and maintenance of remission in inflammatory bowel diseases: a meta-analysis of randomized controlled trials. *Clinical Journal of Gastroenterology* **7,** 1–13 (2014).

89. Ritchie, M. L. & Romanuk, T. N. A Meta-Analysis of Probiotic Efficacy for Gastrointestinal Diseases. *PLoS ONE* **7,** e34938 (2012).

90. Zuccotti, G. *et al.* Probiotics for prevention of atopic diseases in infants: systematic review and meta-analysis. *Allergy* **70,** n/a–n/a (2015).

91. Duan, F. & March, J. C. Engineered bacterial communication prevents Vibrio cholerae virulence in an infant mouse model. *Proceedings of the National Academy of Sciences* **107,** 11260–11264 (June 2010).

92. Hwang, I. Y. *et al.* Engineered probiotic Escherichia coli can eliminate and prevent Pseudomonas aeruginosa gut infection in animal models. *Nature Communications* **8,** 15028 (2017).

93. Lagenaur, L. A. *et al.* Prevention of vaginal SHIV transmission in macaques by a live recombinant Lactobacillus. *Mucosal Immunology* **4,** 648–657 (2011).

94. Steidler, L. *et al.* Treatment of Murine Colitis by Lactococcus lactis Secreting Interleukin-10. *Science* **289,** 1352–1355 (Aug. 2000).

95. Braat, H. *et al.* A Phase I Trial With Transgenic Bacteria Expressing Interleukin-10 in Crohn's Disease. *Clinical Gastroenterology and Hepatology* **4,** 754–759 (June 2006).

96. Takiishi, T. *et al.* Reversal of autoimmune diabetes by restoration of antigen-specific tolerance using genetically modified Lactococcus lactis in mice. *Journal of Clinical Investigation* **122,** 1717–1725 (May 2012).

97. Robert, S. *et al.* Oral delivery of glutamic acid decarboxylase (GAD)-65 and IL10 by lactococcus lactis reverses diabetes in recent-onset NOD mice. *Diabetes* **63,** 2876–2887 (2014).

98. Hamady, Z. Z. R. *et al.* Treatment of colitis with a commensal gut bacterium engineered to secrete human TGF-β1 under the control of dietary xylan. *Inflammatory Bowel Diseases* **17,** 1925–1935 (Sept. 2011).

99. Vandenbroucke, K. *et al.* Orally administered L. lactis secreting an anti-TNF Nanobody demonstrate efficacy in chronic colitis. *Mucosal immunology* **3,** 49–56 (2010).

100. Hamady, Z. Z. R. *et al.* Xylan-regulated delivery of human keratinocyte growth factor-2 to the inflamed colon by the human anaerobic commensal bacterium Bacteroides ovatus. *Gut* **59,** 461–9 (Apr. 2010).

101. Motta, J.-P. *et al.* Food-Grade Bacteria Expressing Elafin Protect Against Inflammation and Restore Colon Homeostasis. *Science Translational Medicine* **4,** 158ra144–158ra144 (2012).

102. Caluwaerts, S. *et al.* AG013, a mouth rinse formulation of Lactococcus lactis secreting human Trefoil Factor 1, provides a safe and efficacious therapeutic tool for treating oral mucositis. *Oral oncology* **46,** 564–70 (2010).

103. Limaye, S. A. *et al.* Phase 1b, multicenter, single blinded, placebo-controlled, sequential dose escalation study to assess the safety and tolerability of topically applied AG013 in subjects with locally advanced head and neck cancer receiving induction chemotherapy. *Cancer* **119,** 4268–4276 (2013).

104. Chen, Z. *et al.* Incorporation of therapeutically modified bacteria into gut microbiota inhibits obesity. *The Journal of clinical investigation* **119,** 1–16 (June 2014).

105. Duan, F. F., Liu, J. H. & March, J. C. Engineered Commensal Bacteria Reprogram Intestinal Cells Into Glucose-Responsive Insulin-Secreting Cells for the Treatment of Diabetes. *Diabetes* **64,** 1794–1803 (May 2015).

106. Baxter, N. T., Ruffin, M. T., Rogers, M. A. & Schloss, P. D. Microbiota-based model improves the sensitivity of fecal immunochemical test for detecting colonic lesions. *Genome Medicine* **8,** 1–10 (2016).

107. Doherty, M. K. *et al.* Fecal Microbiota Signatures Are Associated with Response to Ustekinumab Therapy among Crohn's Disease Patients. *mBio* **9,** e02120–17 (2018).

108. Courbet, A., Endy, D., Renard, E., Molina, F. & Bonnet, J. Detection of pathological biomarkers in human clinical samples via amplifying genetic switches and logic gates. *Science translational medicine* **7,** 289–83 (May 2015).

109. Kotula, J. W. *et al.* Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proceedings of the National Academy of Sciences of the United States of America* **111,** 4838–43 (Apr. 2014).

110. Daeffler, K. N.-M. *et al.* Engineering bacterial thiosulfate and tetrathionate sensors for detecting gut inflammation. *Molecular Systems Biology* **13,** 923 (Apr. 2017).

111. Riglar, D. T. *et al.* Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. *Nature Biotechnology* **35,** 653–658 (2017).

112. Kassam, Z., Lee, C. H., Yuan, Y. & Hunt, R. H. Fecal Microbiota Transplantation for Clostridium difficile Infection: Systematic Review and Meta-Analysis. *The American Journal of Gastroenterology* **108,** 500–508 (2013).

113. Van Nood, E. *et al.* Duodenal Infusion of Donor Feces for Recurrent ¡¿Clostridium difficile¡/¿. *New England Journal of Medicine* **368,** 407–415 (2013).

114. Smith, M. B., Kelly, C. & Alm, E. J. How to regulate faecal transplants. *Nature* **506,** 290–1 (2014).

115. Petrof, E. *et al.* Stool substitute transplant therapy for the eradication of Clostridium difficile infection: 'RePOOPulating' the gut. *Microbiome* **1,** 3 (2013).

116. Ratner, M. Microbial cocktails join fecal transplants in IBD treatment trials. *Nature Biotechnology* **33,** 787–788 (Aug. 2015).

117. Ianiro, G., Bibbò, S., Scaldaferri, F., Gasbarrini, A. & Cammarota, G. Fecal Microbiota Transplantation in Inflammatory Bowel Disease. *Medicine* **93,** e97 (2014).

118. Moayyedi, P. *et al.* Fecal Microbiota Transplantation Induces Remission in Patients With Active Ulcerative Colitis in a Randomized Controlled Trial. *Gastroenterology* **149,** 102–109.e6 (2015).

119. Shen, T.-c. D. *et al.* Engineering the gut microbiota to treat hyperammonemia. *Journal of Clinical Investigation* **125,** 2841–2850 (July 2015).

120. Walsh, C. J. *et al.* In silico identification of bacteriocin gene clusters in the gastrointestinal tract, based on the Human Microbiome Project's reference genome database. *BMC Microbiology* **15,** 183 (2015).

121. Zheng, J., Gänzle, M. G., Lin, X. B., Ruan, L. & Sun, M. Diversity and dynamics of bacteriocins from human microbiome. *Environmental Microbiology* **17,** 2133–2143 (2015).

122. Kommineni, S. *et al.* Bacteriocin production augments niche competition by enterococci in the mammalian gastrointestinal tract. *Nature* (2015).

123. Corr, S. C. *et al.* Bacteriocin production as a mechanism for the antiinfective activity of Lactobacillus salivarius UCC118. *Proceedings of the National Academy of Sciences* **104,** 7617–7621 (2007).

124. Millette, M. *et al.* Capacity of human nisin- and pediocin-producing lactic Acid bacteria to reduce intestinal colonization by vancomycin-resistant enterococci. *Applied and environmental microbiology* **74,** 1997–2003 (2008).

125. D'Herelle, F. Sur un microbe invisible antagoniste des bacilles dysentériques. *CR Acad. Sci. Paris* **165,** 373–375 (1917).

126. Twort, F. W. An Investigation on the Nature of Ultra-Microscopic Viruses. *The Lancet* **186,** 1241–1243 (Dec. 1915).

127. Kingwell, K. Bacteriophage therapies re-enter clinical trials. *Nature Reviews Drug Discovery* **14,** 515–516 (2015).

128. Reardon, S. Phage therapy gets revitalized. *Nature* **510,** 15–16 (2014).

129. Sulakvelidze, A., Alavidze, Z. & Morris, J. G. Bacteriophage Therapy. *Antimicrobial Agents and Chemotherapy* **45,** 649–659 (Mar. 2001).

130. Maura, D. *et al.* Intestinal colonization by enteroaggregative Escherichia coli supports long-term bacteriophage replication in mice. *Environmental Microbiology* **14,** 1844–1854 (2012).

131. Weiss, M. *et al.* In vivo replication of T4 and T7 bacteriophages in germ-free mice colonized with Escherichia coli. *Virology* **393,** 16–23 (Oct. 2009).

132. Mills, S. *et al.* Movers and shakers: Influence of bacteriophages in shaping the mammalian gut microbiota. *Gut Microbes* **4,** 4–16 (2013).

133. Reyes, A. *et al.* Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* **466,** 334–8 (July 2010).

134. Minot, S. *et al.* The human gut virome: inter-individual variation and dynamic response to diet. *Genome research* **21,** 1616–25 (2011).

135. Norman, J. M. *et al.* Disease-Specific Alterations in the Enteric Virome in Inflammatory Bowel Disease. *Cell* **160,** 447–460 (2015).

136. Modi, S. R., Lee, H. H., Spina, C. S. & Collins, J. J. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* **499,** 219–222 (2013).

137. Barr, J. J. *et al.* Subdiffusive motion of bacteriophage in mucosal surfaces increases the frequency of bacterial encounters. *Proceedings of the National Academy of Sciences,* 201508355 (2015).

138. Ando, H., Lemire, S., Pires, D. P. & Lu, T. K. Engineering Modular Viral Scaffolds for Targeted Bacterial Population Editing. *Cell Systems* **1,** 187–196 (2015).

139. Lu, T. K. & Collins, J. J. Dispersing biofilms with engineered enzymatic bacteriophage. *Proceedings of the National Academy of Sciences* **104,** 11197–11202 (July 2007).

140. Edgar, R., Friedman, N., Molshanski-Mor, S. & Qimron, U. Reversing bacterial resistance to antibiotics by phage-mediated delivery of dominant sensitive genes. *Applied and environmental microbiology* **78,** 744–51 (Feb. 2012).

141. Lu, T. K. & Collins, J. J. Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy. *Proceedings of the National Academy of Sciences* **106,** 4629–4634 (Mar. 2009).

142. Hagens, S., Habel, A., von Ahsen, U., von Gabain, A. & Bläsi, U. Therapy of Experimental Pseudomonas Infections with a Nonreplicating Genetically Modified Phage. *Antimicrobial Agents and Chemotherapy* **48,** 3817–3822 (Oct. 2004).

143. Krom, R. J., Bhargava, P., Lobritz, M. a. & Collins, J. J. Engineered Phagemids for Nonlytic, Targeted Antibacterial Therapies. *Nano Letters* **15,** 4808–4813 (2015).

144. Westwater, C. *et al.* Use of Genetically Engineered Phage To Deliver Antimicrobial Agents to Bacteria : an Alternative Therapy for Treatment of Bacterial Infections. *Antimicrobial agents and chemotherapy* **47,** 1301–1307 (Apr. 2003).

145. Bikard, D. *et al.* Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nature biotechnology,* 1–6 (Oct. 2014).

146. Citorik, R. J., Mimee, M. & Lu, T. K. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nature biotechnology* **32**, 1141–5 (Dec. 2014).

147. Kiro, R., Shitrit, D. & Qimron, U. Efficient engineering of a bacteriophage genome using the type I-E CRISPR-Cas system. *RNA Biology* **11**, 42–44 (2014).

148. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth* **6**, 343–345 (May 2009).

149. Centers for Disease Control and Prevention, National Center for Emerging and Zoonotic Infectious Diseases & Division of Healthcare Quality Promotion. *Guidance for Control of Carbapenem-resistant Enterobacteriaceae (CRE) 2012 CRE Toolkit* tech. rep. (2012).

150. Nordmann, P., Dortet, L. & Poirel, L. Carbapenem resistance in Enterobacteriaceae: here is the storm! *Trends in molecular medicine* **18**, 263–72 (May 2012).

151. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science (New York, N.Y.)* **315**, 1709–12 (Mar. 2007).

152. Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (Nov. 2010).

153. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science (New York, N.Y.)* **339**, 823–6 (Feb. 2013).

154. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602–607 (Mar. 2011).

155. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)* **337**, 816–21 (Aug. 2012).

156. Rasheed, J. K. *et al.* Characterization of the extended-spectrum beta-lactamase reference strain, Klebsiella pneumoniae K6 (ATCC 700603), which produces the novel enzyme SHV-18. *Antimicrobial agents and chemotherapy* **44,** 2382–8 (Sept. 2000).

157. Rasheed, J. K. *et al.* New Delhi metallo-β-lactamase-producing Enterobacteriaceae, United States. *Emerging infectious diseases* **19,** 870–8 (June 2013).

158. Bikard, D., Hatoum-Aslan, A., Mucida, D. & Marraffini, L. A. CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell host & microbe* **12,** 177–86 (Aug. 2012).

159. Gomaa, A. A. *et al.* Programmable Removal of Bacterial Strains by Use of Genome-Targeting CRISPR-Cas Systems. *mBio* **5** (Jan. 2013).

160. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic acids research* **25,** 1203–10 (Mar. 1997).

161. Pérez-Mendoza, D. & de la Cruz, F. Escherichia coli genes affecting recipient ability in plasmid conjugation: are there any? *BMC genomics* **10,** 71 (Jan. 2009).

162. Jacoby, G. A. Mechanisms of resistance to quinolones. *Clinical infectious diseases* **41 Suppl 2,** S120–6 (July 2005).

163. Pennington, J. M. & Rosenberg, S. M. Spontaneous DNA breakage in single living Escherichia coli cells. *Nature genetics* **39,** 797–802 (June 2007).

164. Dwyer, D. J., Kohanski, M. A., Hayete, B. & Collins, J. J. Gyrase inhibitors induce an oxidative damage cellular death pathway in Escherichia coli. *Mol Syst Biol* **3** (Mar. 2007).

165. Hayes, F. Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. *Science (New York, N.Y.)* **301,** 1496–9 (Sept. 2003).

166. Mnif, B. *et al.* Molecular characterization of addiction systems of plasmids encoding extended-spectrum beta-lactamases in Escherichia coli. *The Journal of antimicrobial chemotherapy* **65,** 1599–603 (Aug. 2010).

167. Kaper, J. B., Nataro, J. P. & Mobley, H. L. Pathogenic Escherichia coli. *Nature reviews. Microbiology* **2,** 123–40 (Feb. 2004).

168. Desbois, A. P. & Coote, P. J. Wax moth larva (Galleria mellonella): an in vivo model for assessing the efficacy of antistaphylococcal agents. *The Journal of antimicrobial chemotherapy* **66,** 1785–90 (Aug. 2011).

169. Sonnenburg, J. L. & Fischbach, M. A. Community health care: therapeutic opportunities in the human microbiome. *Science Translational Medicine* **3,** 78ps12 (Apr. 2011).

170. Paddon, C. J. *et al.* High-level semi-synthetic production of the potent anti-malarial artemisinin. *Nature* **496,** 528–32 (Apr. 2013).

171. Duan, F. & March, J. C. Engineered bacterial communication prevents Vibrio cholerae virulence in an infant mouse model. *Proceedings of the National Academy of Sciences of the United States of America* **107,** 11260–4 (June 2010).

172. Lu, T. K. & Collins, J. J. Dispersing biofilms with engineered enzymatic bacteriophage. *Proceedings of the National Academy of Sciences of the United States of America* **104,** 11197–202 (July 2007).

173. Lu, T. K. & Collins, J. J. Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy. *Proceedings of the National Academy of Sciences of the United States of America* **106,** 4629–34 (Mar. 2009).

174. Edgar, R., Friedman, N., Molshanski-Mor, S. & Qimron, U. Reversing bacterial resistance to antibiotics by phage-mediated delivery of dominant sensitive genes. *Applied and environmental microbiology* **78,** 744–51 (Feb. 2012).

175. Seed, K. D., Lazinski, D. W., Calderwood, S. B. & Camilli, A. A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature* **494,** 489–91 (Feb. 2013).

176. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature biotechnology* **31**, 233–9 (Mar. 2013).

177. Vercoe, R. B. *et al.* Cytotoxic Chromosomal Targeting by CRISPR/Cas Systems Can Reshape Bacterial Genomes and Expel or Remodel Pathogenicity Islands. *PLoS Genet* **9**, e1003454 (Apr. 2013).

178. Williams, J. J. & Hergenrother, P. J. Artificial activation of toxin-antitoxin systems as an antibacterial strategy. *Trends in microbiology* **20**, 291–8 (June 2012).

179. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nature reviews. Microbiology* **8**, 317–27 (May 2010).

180. Lin, T.-Y. *et al.* A T3 and T7 recombinant phage acquires efficient adsorption and a broader host range. *PloS one* **7**, e30954 (Jan. 2012).

181. Abedon, S. T., Kuhl, S. J., Blasdel, B. G. & Kutter, E. M. Phage treatment of human infections. *Bacteriophage* **1**, 66–85 (Jan. 2011).

182. Bruttin, A. & Brüssow, H. Human volunteers receiving Escherichia coli phage T4 orally: a safety test of phage therapy. *Antimicrobial agents and chemotherapy* **49**, 2874–8 (July 2005).

183. Merril, C. R. *et al.* Long-circulating bacteriophage as antibacterial agents. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 3188–92 (Apr. 1996).

184. Hagens, S. & Bläsi, U. Genetically modified filamentous phage as bactericidal agents: a pilot study. *Letters in applied microbiology* **37**, 318–23 (Jan. 2003).

185. Saeidi, N. *et al.* Engineering microbes to sense and eradicate Pseudomonas aeruginosa, a human pathogen. *Molecular systems biology* **7**, 521 (Jan. 2011).

186. Esvelt, K. M., Smidler, A. L., Catteruccia, F. & Church, G. M. Concerning RNA-guided gene drives for the alteration of wild populations. *eLife*, e03401 (July 2014).

187. Datta, S., Costantino, N. & Court, D. L. A set of recombineering plasmids for gram-negative bacteria. *Gene* **379,** 109–15 (Sept. 2006).

188. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods* **6,** 343–5 (May 2009).

189. Sikorski, R. S. & Hieter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics* **122,** 19–27 (May 1989).

190. Clinical and Laboratory Standards Institute. *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically; Approved Standard–Seventh Edition* 64 (2006).

191. Chung, C. T., Niemela, S. L. & Miller, R. H. One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. *Proceedings of the National Academy of Sciences of the United States of America* **86,** 2172–5 (Apr. 1989).

192. Chasteen, L., Ayriss, J., Pavlik, P. & Bradbury, A. R. M. Eliminating helper phage from phage display. *Nucleic acids research* **34,** e145 (Jan. 2006).

193. Westwater, C. *et al.* Use of genetically engineered phage to deliver antimicrobial agents to bacteria: an alternative therapy for treatment of bacterial infections. *Antimicrobial agents and chemotherapy* **47,** 1301–7 (Apr. 2003).

194. Dong, D., Sutaria, S., Hwangbo, J. Y. & Chen, P. A simple and rapid method to isolate purer M13 phage by isoelectric precipitation. *Applied microbiology and biotechnology* **97,** 8023–9 (Sept. 2013).

195. Ramarao, N., Nielsen-Leroux, C. & Lereclus, D. The insect Galleria mellonella as a powerful infection model to investigate bacterial pathogenesis. *Journal of visualized experiments : JoVE,* e4392 (Jan. 2012).

196. Mimee, M., Tucker, A. C., Voigt, C. A. & Lu, T. K. Programming a Human Commensal Bacterium, Bacteroides thetaiotaomicron, to Sense and Respond to Stimuli in the Murine Gut Microbiota. *Cell Systems,* 1–10 (2015).

197. Salyers, A. A. Bacteroides of the Human Lower Intestinal Tract, 293–313 (1984).

198. Lee, S. M. *et al.* Bacterial colonization factors control specificity and stability of the gut microbiota. *Nature* **501**, 426–9 (Sept. 2013).

199. Martens, E. C., Chiang, H. C. & Gordon, J. I. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell host & microbe* **4**, 447–57 (Nov. 2008).

200. Nemunaitis, J. *et al.* Pilot trial of genetically modified, attenuated Salmonella expressing the E. coli cytosine deaminase gene in refractory cancer patients. *Cancer gene therapy* **10**, 737–744 (2003).

201. Rothman, J. & Paterson, Y. Live-attenuated Listeria-based immunotherapy. *Expert Review of Vaccines* **12**, 493–504 (2013).

202. Steidler, L. *et al.* Biological containment of genetically modified Lactococcus lactis for intestinal delivery of human interleukin 10. *Nature biotechnology* **21**, 785–789 (2003).

203. Leavitt, J. M. & Alper, H. S. Advances and current limitations in transcript-level control of gene expression. *Current Opinion in Biotechnology* **34**, 98–104 (2015).

204. Nielsen, A. A., Segall-Shapiro, T. H. & Voigt, C. a. Advances in genetic circuit design: novel biochemistries, deep part mining, and precision gene expression. *Current opinion in chemical biology* **17**, 878–92 (Dec. 2013).

205. Brophy, J. A. N. & Voigt, C. A. Principles of genetic circuit design. *Nature Methods* **11**, 508–520 (Apr. 2014).

206. Geyer, P. K. The role of insulator elements in defining domains of gene expression. *Current Opinion in Genetics and Development* **7**, 242–248 (1997).

207. Hamady, Z. Z. R. *et al.* Identification and use of the putative Bacteroides ovatus xylanase promoter for the inducible production of recombinant human proteins. *Microbiology (Reading, England)* **154**, 3165–74 (Oct. 2008).

208. Parker, A. C. & Smith, J. C. Development of an IPTG inducible expression vector adapted for Bacteroides fragilis. *Plasmid* **68**, 86–92 (Sept. 2012).

209. Smith, C. J., Rogers, M. B. & McKee, M. L. Heterologous gene expression in Bacteroides fragilis. *Plasmid* **27**, 141–54 (Mar. 1992).

210. Wang, J., Shoemaker, N. B., Wang, G.-R. & Salyers, A. A. Characterization of a Bacteroides Mobilizable Transposon, NBU2, Which Carries a Functional Lincomycin Resistance Gene. *Journal of Bacteriology* **182**, 3559–3571 (June 2000).

211. Bayley, D. P., Rocha, E. R. & Smith, C. J. Analysis of cepA and other Bacteroides fragilis genes reveals a unique promoter structure. *FEMS microbiology letters* **193**, 149–54 (Dec. 2000).

212. Vingadassalom, D. *et al.* An unusual primary sigma factor in the Bacteroidetes phylum. *Molecular Microbiology* **56**, 888–902 (2005).

213. Wegmann, U., Horn, N. & Carding, S. R. Defining the bacteroides ribosomal binding site. *Applied and environmental microbiology* **79**, 1980–9 (Mar. 2013).

214. Accetto, T. & Avguštin, G. Inability of Prevotella bryantii to form a functional Shine-Dalgarno interaction reflects unique evolution of ribosome binding sites in Bacteroidetes. *PloS one* **6**, e22914 (Jan. 2011).

215. Coyne, M. J., Weinacht, K. G., Krinos, C. M. & Comstock, L. E. Mpi recombinase globally modulates the surface architecture of a human commensal bacterium. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 10446–10451 (2003).

216. Bonnet, J., Yin, P., Ortiz, M. E., Subsoontorn, P. & Endy, D. Amplifying genetic logic gates. *Science (New York, N.Y.)* **340**, 599–603 (Mar. 2013).

217. Friedland, A. E. *et al.* Synthetic Gene Networks That Count. *Science* **324**, 1199–1202 (May 2009).

218. Siuti, P., Yazbek, J. & Lu, T. K. Synthetic circuits integrating logic and memory in living cells. *Nature biotechnology* **31**, 448–52 (May 2013).

219. Yang, L. *et al.* Permanent genetic memory with ¿1-byte capacity. *Nature Methods* **11**, 1261–6 (2014).

220. Farzadfard, F. & Lu, T. K. Genomically encoded analog memory with precise in vivo DNA writing in living cell populations. *Science* **346**, 6211 (2014).

221. Mali, P., Esvelt, K. M. & Church, G. M. Cas9 as a versatile tool for engineering biology. *Nat Methods* **10**, 957–963 (2013).

222. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–83 (Feb. 2013).

223. Nielsen, A. A. & Voigt, C. a. Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Molecular systems biology* **10**, 763 (Jan. 2014).

224. Li, L. Y., Shoemaker, N. B. & Salyers, A. A. Characterization of the mobilization region of a Bacteroides insertion element (NBU1) that is excised and transferred by Bacteroides conjugative transposons. *Journal of Bacteriology* **175**, 6588–6598 (1993).

225. Shoemaker, N. B., Wang, G. R. & Salyers, A. a. The Bacteroides mobilizable insertion element, NBU1, integrates into the 3' end of a leu-tRNA gene and has an integrase that is a member of the lambda integrase family. *Journal of Bacteriology* **178**, 3594–3600 (1996).

226. Rajeev, L., Salyers, A. a. & Gardner, J. F. Characterization of the integrase of NBU1, a Bacteroides mobilizable transposon. *Molecular Microbiology* **61**, 978–990 (Aug. 2006).

227. Goto, T., Tanaka, K., Minh Tran, C. & Watanabe, K. Complete sequence of {pBFUK1}, a carbapenemase-harboring mobilizable plasmid from Bacteroides fragilis, and distribution of {pBFUK1-like} plasmids among carbapenem-resistant B. fragilis clinical isolates. *The Journal of antibiotics* **66**, 239–242 (2013).

228. Parker, A. C. & Smith, C. J. Genetic and biochemical analysis of a novel ambler class A β-lactamase responsible for cefoxitin resistance in Bacteroides species. *Antimicrobial Agents and Chemotherapy* **37**, 1028–1036 (1993).

229. Rogers, M. B., Bennett, T. K., Payne, C. M. & Smith, C. J. Insertional activation of cepA leads to high-level β-lactamase expression in Bacteroides fragilis clinical isolates. *Journal of Bacteriology* **176**, 4376–4384 (1994).

230. Boni, I. V., Isaeva, D. M., Musychenko, M. L. & Tzareva, N. V. Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic acids research* **19**, 155–62 (Jan. 1991).

231. Patel, E. H., Paul, L. V., Patrick, S. & Abratt, V. R. Rhamnose catabolism in Bacteroides thetaiotaomicron is controlled by the positive transcriptional regulator RhaR. *Research in microbiology* **159**, 678–84 (2008).

232. Sonnenburg, E. D. *et al.* A hybrid two-component system protein of a prominent human gut symbiont couples glycan sensing in vivo to carbohydrate metabolism. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 8834–8839 (2006).

233. Martens, E. C. *et al.* Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. *PLoS biology* **9**, e1001221 (Dec. 2011).

234. Grindley, N. D. F., Whiteson, K. L. & Rice, P. a. Mechanisms of site-specific recombination. *Annual review of biochemistry* **75**, 567–605 (2006).

235. Cullen, T. W. *et al.* Antimicrobial peptide resistance mediates resilience of prominent gut commensals during inflammation. *Science* **347**, 170–175 (Jan. 2015).

236. Sonnenburg, E. D. *et al.* Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell* **141**, 1241–52 (June 2010).

237. Lee, J.-Y. *et al.* Contribution of the 7β-hydroxysteroid dehydrogenase from Ruminococcus gnavus N53 to ursodeoxycholic acid formation in the human colon. *Journal of lipid research* **54,** 3062–9 (2013).

238. Bloom, S. M. *et al.* Commensal Bacteroides species induce colitis in host-genotype-specific fashion in a mouse model of inflammatory bowel disease. *Cell Host and Microbe* **9,** 390–403 (2011).

239. Cox, R. S., Surette, M. G. & Elowitz, M. B. Programming gene expression with combinatorial promoters. *Molecular systems biology* **3,** 145 (Jan. 2007).

240. Kosuri, S. *et al.* Composability of regulatory sequences controlling transcription and translation in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **110,** 14024–9 (2013).

241. Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nature biotechnology* **27,** 946–50 (Oct. 2009).

242. Jensen, P. R. & Hammer, K. The Sequence of Spacers between the Consensus Sequences Modulates the Strength of Prokaryotic Promoters The Sequence of Spacers between the Consensus Sequences Modulates the Strength of Prokaryotic Promoters. *Applied and environmental microbiology* **64** (1998).

243. Kleerebezem, M., Beerthuyzen, M. M., Vaughan, E. E., de Vos, W. M. & Kuipers, O. P. Controlled gene expression systems for lactic acid bacteria: transferable nisin-inducible expression cassettes for Lactococcus, Leuconostoc, and Lactobacillus spp. *Applied and environmental microbiology* **63,** 4581–4 (1997).

244. Liu, C. H., Lee, S. M., Vanlare, J. M., Kasper, D. L. & Mazmanian, S. K. Regulation of surface architecture by symbiotic bacteria mediates host colonization. *Proceedings of the National Academy of Sciences of the United States of America* **105,** 3951–6 (Mar. 2008).

245. Claesen, J. & Fischbach, M. a. Synthetic Microbes As Drug Delivery Systems. *ACS synthetic biology,* DOI: 10.1021/sb500258b (Aug. 2014).

246. Mahowald, M. a. *et al.* Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5859–64 (Apr. 2009).

247. Bacic, M. K. & Smith, C. J. Laboratory maintenance and cultivation of bacteroides species. *Current protocols in microbiology* **Chapter 13**, Unit 13C.1 (May 2008).

248. Simon, R., Priefer, U. & Pühler, A. A broad host range mobilization system for in vivo genetic engineering: Transposon mutagenesis in Gram negative bacteria. *Nature biotechnology* **1**, 784–791 (Nov. 1983).

249. Kessler, B., De Lorenzo, V. & Timmis, K. N. A general system to integrate lacZ fusions into the chromosomes of gram-negative eubacteria: Regulation of the Pm promoter of the TOL plasmid studied with all controlling elements in monocopy. *Molecular and General Genetics* **233**, 293–301 (1992).

250. Koropatkin, N. M., Cameron, E. a. & Martens, E. C. How glycan metabolism shapes the human gut microbiota. *Nature reviews. Microbiology* **10**, 323–35 (May 2012).

251. Koropatkin, N. M., Martens, E. C., Gordon, J. I. & Smith, T. J. Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. *Structure (London, England : 1993)* **16**, 1105–15 (July 2008).

252. Hall, M. P. *et al.* Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate. *ACS Chemical Biology* **7**, 1848–1857 (Nov. 2012).

253. Cubitt, A. B. *et al.* Understanding, improving and using green fluorescent proteins. *Trends in biochemical sciences* **20**, 448–55 (1995).

254. Lobo, L. a., Smith, C. J. & Rocha, E. R. Flavin mononucleotide (FMN)-based fluorescent protein (FbFP) as reporter for gene expression in the anaerobe Bacteroides fragilis. *FEMS microbiology letters* **317**, 67–74 (Apr. 2011).

255. Mastropaolo, M. D., Thorson, M. L. & Stevens, A. M. Comparison of Bacteroides thetaiotaomicron and Escherichia coli 16S rRNA gene expression signals. *Microbiology (Reading, England)* **155**, 2683–93 (Aug. 2009).

256. Radeck, J. *et al.* The Bacillus BioBrick Box: generation and evaluation of essential genetic building blocks for standardized work with Bacillus subtilis. *Journal of biological engineering* **7**, 29 (2013).

257. Waidmann, M. S., Bleichrodt, F. S., Laslo, T. & Riedel, C. U. Bacterial luciferase reporters: the Swiss army knife of molecular biology. *Bioengineered bugs* **2**, 8–16 (2011).

258. CLSI. *Performance Standards for Antimicrobial Susceptibility Testing; Seventeenth Informational Supplement* tech. rep. (CLSI, 2007), 1–180.

259. Giraud, M. F. & Naismith, J. H. The rhamnose pathway. *Current Opinion in Structural Biology* **10**, 687–696 (2000).

260. Sender, R., Fuchs, S. & Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biology* **14**, 1–14 (2016).

261. Neff, C. P. *et al.* Diverse Intestinal Bacteria Contain Putative Zwitterionic Capsular Polysaccharides with Anti-inflammatory Properties. *Cell Host & Microbe,* 1–13 (2016).

262. Kulp, A. & Kuehn, M. J. Biological functions and biogenesis of secreted bacterial outer membrane vesicles. *Annual review of microbiology* **64**, 163–184 (2010).

263. Schwechheimer, C. & Kuehn, M. J. Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. *Nature reviews. Microbiology* **13**, 605–19 (2015).

264. Schwechheimer, C., Sullivan, C. J. & Kuehn, M. J. Envelope control of outer membrane vesicle production in Gram-negative bacteria. *Biochemistry* **52**, 3031–3040 (2013).

265. McBroom, A. J. & Kuehn, M. J. Release of outer membrane vesicles by Gram-negative bacteria is a novel envelope stress response. *Molecular Microbiology* **63**, 545–558 (2007).

266. Ellis, T. N. & Kuehn, M. J. Virulence and immunomodulatory roles of bacterial outer membrane vesicles. *Microbiology and molecular biology reviews : MMBR* **74**, 81–94 (2010).

267. Kaparakis-Liaskos, M. & Ferrero, R. L. Immune modulation by bacterial outer membrane vesicles. *Nature Reviews Immunology* **15**, 375–387 (2015).

268. Shen, Y. *et al.* Outer membrane vesicles of a human commensal mediate immune regulation and disease protection. *Cell host & microbe* **12**, 509–20 (Oct. 2012).

269. Chu, H. *et al.* Gene-microbiota interactions contribute to the pathogenesis of inflammatory bowel disease. *Science* **9948**, 1–10 (2016).

270. Hickey, C. A. *et al.* Colitogenic Bacteroides thetaiotaomicron Antigens Access Host Immune Cells in a Sulfatase-Dependent Manner via Outer Membrane Vesicles. *Cell Host & Microbe* **17**, 672–680 (2015).

271. Chen, D. J. *et al.* Delivery of foreign antigens by engineered outer membrane vesicle vaccines. *Proceedings of the National Academy of Sciences* **107**, 3099–3104 (Feb. 2010).

272. Kesty, N. C. & Kuehn, M. J. Incorporation of heterologous outer membrane and periplasmic proteins into Escherichia coli outer membrane vesicles. *The Journal of biological chemistry* **279**, 2069–76 (Jan. 2004).

273. Round, J. L. *et al.* The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. *Science* **332**, 974–7 (2011).

274. Schwechheimer, C., Rodriguez, D. L. & Kuehn, M. J. NlpI-mediated modulation of outer membrane vesicle production through peptidoglycan dynamics in Escherichia coli. *Microbiology* **4**, 375–389 (2015).

275. Elhenawy, W., Debelyy, M. O. & Feldman, M. F. Preferential Packing of Acidic Glycosidases and Proteases into Bacteroides Outer Membrane Vesicles. *mBio* **5,** e00909–14–e00909–14 (Mar. 2014).

276. Buddelmeijer, N. The molecular mechanism of bacterial lipoprotein modification-How, when and why? *FEMS Microbiology Reviews* **39,** 246–261 (2015).

277. Zückert, W. R. Secretion of Bacterial Lipoproteins: Through the Cytoplasmic Membrane, the Periplasm and Beyond. *Biochimica et Biophysica Acta - Molecular Cell Research* **1843,** 1509–1516 (2014).

278. Rahman, O., Cummings, S. P., Harrington, D. J. & Sutcliffe, I. C. Methods for the bioinformatic identification of bacterial lipoproteins encoded in the genomes of Gram-positive bacteria. *World Journal of Microbiology and Biotechnology* **24,** 2377–2382 (2008).

279. Nielsen, H. in *Protein Function Prediction: Methods and Protocols* (ed Kihara, D.) 59–73 (Springer New York, New York, NY, 2017).

280. Narita, S. I. & Tokuda, H. Amino acids at positions 3 and 4 determine the membrane specificity of Pseudomonas aeruginosa lipoproteins. *Journal of Biological Chemistry* **282,** 13372–13378 (2007).

281. Yamaguchi, K., Yu, F. & Inouye, M. A single amino acid determinant of the membrane localization of lipoproteins in ¡¿E. coli¡/¿. *Cell* **53,** 423–432 (1988).

282. Murray, P. J. Restraint of inflammatory signaling by interdependent strata of negative regulatory pathways. *Nature Immunology* **13,** 916–924 (Sept. 2012).

283. Glocker, E. O., Kotlarz, D., Klein, C., Shah, N. & Grimbacher, B. IL-10 and IL-10 receptor defects in humans. *Annals of the New York Academy of Sciences* **1246,** 102–107 (2011).

284. Li, L.-J. Role of interleukin-22 in inflammatory bowel disease. *World Journal of Gastroenterology* **20,** 18177 (2014).

285. Sabat, R., Ouyang, W. & Wolk, K. Therapeutic opportunities of the IL-22-IL-22R1 system. *Nature Reviews Drug Discovery* **13,** 21–38 (Dec. 2014).

286. Sugimoto, K. & Ogawa, A. IL-22 ameliorates intestinal inflammation in a mouse model of ulcerative colitis. *The Journal of clinical ...* **118** (2008).

287. Dumoutier, L., de Meester, C., Tavernier, J. & Renauld, J.-C. New Activation Modus of STAT3: A Tyrosine-Less Region of the Interleukin-22 Receptor Recruits STAT3 by Interacting with its Coiled-Coil Domain. *Journal of Biological Chemistry* **284,** 26377–26384 (2009).

288. Dutton, R. J., Boyd, D., Berkmen, M. & Beckwith, J. Bacterial species exhibit diversity in their mechanisms and capacity for protein disulfide bond formation. *Proceedings of the National Academy of Sciences of the United States of America* **105,** 11933–11938 (2008).

289. Chen, X., Zaro, J. L. & Shen, W. C. Fusion protein linkers: Property, design and functionality. *Advanced Drug Delivery Reviews* **65,** 1357–1369 (2013).

290. Hatahet, F., Boyd, D. & Beckwith, J. Disulfide bond formation in prokaryotes: History, diversity and design. *Biochimica et Biophysica Acta - Proteins and Proteomics* **1844,** 1402–1414 (2014).

291. Chen, L. *et al.* Outer membrane vesicles displaying engineered glycotopes elicit protective antibodies. *Proceedings of the National Academy of Sciences,* 201518311 (2016).

292. Hadis, U. *et al.* Intestinal Tolerance Requires Gut Homing and Expansion of FoxP3+ Regulatory T Cells in the Lamina Propria. *Immunity* **34,** 237–246 (2011).

293. Kim, K. S. *et al.* Dietary antigens limit mucosal immunity by inducing regulatory T cells in the small intestine. *Science* **5560** (2016).

294. Worbs, T. *et al.* Oral tolerance originates in the intestinal immune system and relics on antigen carriage by dendritic cells. *The Journal of Experimental Medicine* **203,** 519–527 (Mar. 2006).

295. Malo, M. S. *et al.* Intestinal alkaline phosphatase promotes gut bacterial growth by reducing the concentration of luminal nucleotide triphosphates. *AJP: Gastrointestinal and Liver Physiology* **306,** G826–G838 (2014).

296. Ramasamy, S. *et al.* Intestinal alkaline phosphatase has beneficial effects in mouse models of chronic colitis. *Inflammatory Bowel Diseases* **17,** 532–542 (Feb. 2011).

297. Yeste, A. *et al.* Tolerogenic nanoparticles inhibit T cell–mediated autoimmunity through SOCS2. *Science Signaling* **9,** 1–12 (2016).

298. Pishesha, N. *et al.* Engineered erythrocytes covalently linked to antigenic peptides can protect against autoimmune disease. *Proceedings of the National Academy of Sciences,* 201701746 (2017).

299. Perdicchio, M. *et al.* Sialic acid-modified antigens impose tolerance via inhibition of T-cell proliferation and de novo induction of regulatory T cells. *Proceedings of the National Academy of Sciences* **113,** 201507706 (2016).

300. Brandt, E. B. *et al.* Mast cells are required for experimental oral allergen – induced diarrhea. *Journal of clinical investigation* **112,** 1666–77 (2003).

301. Slomovic, S., Pardee, K. & Collins, J. J. Synthetic biology devices for in vitro and in vivo diagnostics. *Proceedings of the National Academy of Sciences* **112,** 14429–14435 (2015).

302. Roggo, C. & van der Meer, J. R. Miniaturized and integrated whole cell living bacterial sensors in field applicable autonomous devices. *Current Opinion in Biotechnology* **45,** 24–33 (2017).

303. Lim, B., Zimmermann, M., Barry, N. A. & Goodman, A. L. Engineered Regulatory Systems Modulate Gene Expression of Human Commensals in the Gut. *Cell* **169,** 547–558.e15 (2017).

304. Pickard, J. M. *et al.* Rapid fucosylation of intestinal epithelium sustains host–commensal symbiosis in sickness. *Nature* **514,** 638–641 (2014).

305. Borkowski, O., Ceroni, F., Stan, G. B. & Ellis, T. Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Current Opinion in Microbiology* **33,** 123–130 (2016).

306. Otis, B. & Parviz, B. *Introducing our smart contact lens project* 2014.

307. Wang, H. Magnetic sensors for diagnostic medicine: CMOS-based magnetic particle detectors for medical diagnosis applications. *IEEE Microwave Magazine* **14,** 110–130 (2013).

308. Norian, H., Field, R. M., Kymissis, I. & Shepard, K. L. An integrated CMOS quantitative-polymerase-chain-reaction lab-on-chip for point-of-care diagnostics. *Lab Chip* **14,** 4076–4084 (2014).

309. Iddan, G., Meron, G., Glukhovsky, A. & Swain, P. Wireless capsule endoscopy. *Nature* **405,** 417–417 (May 2000).

310. Van der Schaar, P. J. *et al.* A novel ingestible electronic drug delivery and monitoring device. *Gastrointestinal Endoscopy* **78,** 520–528 (2013).

311. Kalantar-Zadeh, K. *et al.* A human pilot trial of ingestible electronic capsules capable of sensing different gases in the gut. *Nature Electronics 2017 1:1* **1,** 79 (2018).

312. Hafezi, H. *et al.* An ingestible sensor for measuring medication adherence. *IEEE Transactions on Biomedical Engineering* **62,** 99–109 (2015).

313. Nadeau, P., Mimee, M., Carim, S., Lu, T. K. & Chandrakasan, A. P. Nanowatt Circuit Interface to Whole-Cell Bacterial Sensors. *ISSCC,* 352–354 (2017).

314. Hearnshaw, S. A. *et al.* Acute upper gastrointestinal bleeding in the UK: patient characteristics, diagnoses and outcomes in the 2007 UK audit. *Gut* **60,** 1327–1335 (2011).

315. Rockey, D. C., Koch, J., Cello, J. P., Sanders, L. L. & McQuaid, K. Relative frequency of upper gastrointestinal and colonic lesions in patients with positive fecal occult-blood tests. *The New England journal of medicine* **339,** 153–9 (July 1998).

316. Barkun, A., Bardou, M. & Marshall, J. K. Clinical Guidelines Consensus Recommendations for Managing Patients with Nonvariceal Upper Gastrointestinal Bleeding. *Annals of Internal Medicine* **139**, 843–857 (2003).

317. Lee, J. G. *et al.* Endoscopy-based triage significantly reduces hospitalization rates and costs of treating upper GI bleeding: a randomized controlled trial. *Gastrointestinal endoscopy* **50**, 755–61 (Dec. 1999).

318. Cheng, C.-L. *et al.* Predictors of rebleeding and mortality in patients with high-risk bleeding peptic ulcers. *Digestive diseases and sciences* **55**, 2577–83 (2010).

319. Lechardeur, D. *et al.* Discovery of intracellular heme-binding protein HrtR, which controls heme efflux by the conserved HrtB-HrtA transporter in Lactococcus lactis. *The Journal of biological chemistry* **287**, 4752–8 (Feb. 2012).

320. Sawai, H., Yamanaka, M., Sugimoto, H., Shiro, Y. & Aono, S. Structural basis for the transcriptional regulation of heme homeostasis in Lactococcus lactis. *The Journal of biological chemistry* **287**, 30755–68 (Aug. 2012).

321. Close, D. *et al.* The evolution of the bacterial luciferase gene cassette (lux) as a real-time bioreporter. *Sensors* **12**, 732–752 (2012).

322. Torres, a. G. & Payne, S. M. Haem iron-transport system in enterohaemorrhagic Escherichia coli O157:H7. *Molecular microbiology* **23**, 825–33 (Feb. 1997).

323. Nobles, C. L., Clark, J. R., Green, S. I. & Maresso, A. W. A dual component heme biosensor that integrates heme transport and synthesis in bacteria. *Journal of Microbiological Methods* **118**, 7–17 (2015).

324. Espah Borujeni, A., Channarasappa, A. S. & Salis, H. M. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Research* **42**, 2646–2659 (Feb. 2014).

325. Lanas, A. & Chan, F. K. L. Peptic ulcer disease. *The Lancet* **390**, 613–624 (2017).

326. Eltoukhy, H., Salama, K. & Gamal, A. E. A 0.18-um CMOS Bioluminescence Detection Lab-on-Chip. *IEEE Journal of Solid-State Circuits* **41**, 651–661 (2006).

327. Singh, R. R., Leng, L., Guenther, A. & Genov, R. A CMOS-microfluidic chemiluminescence contact imaging microsystem. *IEEE Journal of Solid-State Circuits* **47**, 2822–2833 (2012).

328. Nadeau, P. *et al.* Prolonged energy harvesting for ingestible devices. *Nature Biomedical Engineering* **1**, 0022 (2017).

329. Schuster, M., Joseph Sexton, D., Diggle, S. P. & Peter Greenberg, E. Acyl-Homoserine Lactone Quorum Sensing: From Evolution to Application. *Annual Review of Microbiology* **67**, 43–63 (2013).

330. Balagaddé, F. K. *et al.* A synthetic Escherichia coli predator-prey ecosystem. *Molecular Systems Biology* **4**, 1–8 (2008).

331. Bellinger, A. M. *et al.* Oral, ultra–long-lasting drug delivery: Application toward malaria elimination goals. *Science Translational Medicine* **8**, 365ra157–365ra157 (Nov. 2016).

332. Kirtane, A. R. *et al.* Development of an oral once-weekly drug delivery system for HIV antiretroviral therapy. *Nature Communications* **9** (2018).

333. Kearse, M. *et al.* Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics (Oxford, England)* **28**, 1647–9 (June 2012).

334. Whitaker, W. R., Shepherd, E. S. & Sonnenburg, J. L. Tunable Expression Tools Enable Single-Cell Strain Distinction in the Gut Microbiome. *Cell* **169**, 538–546.e12 (2017).

335. Westwater, C., Schofield, D. A., Schmidt, M. G., Norris, J. S. & Dolan, J. W. Development of a P1 phagemid system for the delivery of DNA into Gram-negative bacteria. *Microbiology* **148**, 943–950 (Apr. 2002).

336. Kilcher, S., Studer, P., Muessner, C., Klumpp, J. & Loessner, M. J. Cross-genus rebooting of custom-made, synthetic bacteriophage genomes in L-form bacteria. *Proceedings of the National Academy of Sciences*, 201714658 (2018).

337. Pawluk, A., Davidson, A. R. & Maxwell, K. L. Anti-CRISPR: discovery, mechanism and function. *Nature reviews. Microbiology* **16**, 12–17 (2018).

338. Shabbir, M. A. *et al.* Bacteria vs. bacteriophages: Parallel evolution of immune arsenals. *Frontiers in Microbiology* **7**, 1–8 (2016).

339. Roach, D. R. *et al.* Synergy between the Host Immune System and Bacteriophage Is Essential for Successful Phage Therapy against an Acute Respiratory Pathogen. *Cell Host & Microbe* **22**, 38–47.e4 (2017).

340. Greenblum, S., Chiu, H.-C., Levy, R., Carr, R. & Borenstein, E. Towards a predictive systems-level model of the human microbiome: progress, challenges, and opportunities. *Current Opinion in Biotechnology* **24**, 810–820 (2013).

341. Levy, R. & Borenstein, E. Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 12804–9 (2013).

342. Donaldson, G. P., Lee, S. M. & Mazmanian, S. K. Gut biogeography of the bacterial microbiota. *Nature Reviews Microbiology* **14**, 20–32 (2015).

343. Earle, K. A. *et al.* Quantitative Imaging of Gut Microbiota Spatial Organization. *Cell Host & Microbe* **18**, 478–488 (2015).

344. Li, H. *et al.* The outer mucus layer hosts a distinct intestinal microbial niche. *Nature Communications* (2015).

345. Nava, G. M., Friedrichsen, H. J. & Stappenbeck, T. S. Spatial organization of intestinal microbiota in the mouse ascending colon. *The ISME journal* **5,** 627–38 (Apr. 2011).

346. Sokol, H. *et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proceedings of the National Academy of Sciences of the United States of America* **105,** 16731–6 (2008).

347. Gevers, D. *et al.* The Treatment-Naive Microbiome in New-Onset Crohn's Disease. *Cell Host & Microbe* **15,** 382–392 (2014).

348. Auchtung, J. M., Robinson, C. D. & Britton, R. A. Cultivation of stable, reproducible microbial communities from different fecal donors using minibioreactor arrays (MBRAs). *Microbiome* **3,** 42 (2015).

349. McDonald, J. A. *et al.* Evaluation of microbial community reproducibility, stability and composition in a human distal gut chemostat model. *Journal of Microbiological Methods* **95,** 167–174 (2013).

350. Van den Abbeele, P. *et al.* Butyrate-producing Clostridium cluster XIVa species specifically colonize mucins in an in vitro gut model. *The ISME journal* **7,** 949–61 (2013).

351. Lukovac, S. *et al.* Differential Modulation by Akkermansia muciniphila and Faecalibacterium prausnitzii of Host Peripheral Lipid Metabolism and Histone Acetylation in Mouse Gut Organoids. *mBio* **5,** e01438–14–e01438–14 (Aug. 2014).

352. Costello, C. M. *et al.* 3D Intestinal Scaffolds for Evaluating the Therapeutic Potential of Probiotics. *Mol. Pharmaceutics,* 2030–2039 (2015).

353. Kim, H. J., Li, H., Collins, J. J. & Ingber, D. E. Contributions of microbiome and mechanical deformation to intestinal bacterial overgrowth and inflammation in a human gut-on-a-chip. *Proceedings of the National Academy of Sciences,* 201522193 (2015).

354. Ceroni, F., Algar, R., Stan, G.-B. & Ellis, T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nature Methods* **12,** 415–418 (2015).

355. Gladstone, E. G., Molineux, I. J. & Bull, J. J. Evolutionary principles and synthetic biology: avoiding a molecular tragedy of the commons with an engineered phage. *Journal of biological engineering* **6,** 13 (Jan. 2012).

356. Springman, R., Molineux, I. J., Duong, C., Bull, R. J. & Bull, J. J. Evolutionary stability of a refactored phage genome. *ACS synthetic biology* **1,** 425–30 (Sept. 2012).

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

358. Danino, T. *et al.* Programmable probiotics for detection of cancer in urine. *Science Translational Medicine* **7,** 289ra84–289ra84 (2015).

359. Ellefson, J. W. *et al.* Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nature Biotechnology* **32,** 97–101 (Nov. 2013).

360. Esvelt, K. M., Carlson, J. C. & Liu, D. R. A system for the continuous directed evolution of biomolecules. *Nature* **472,** 499–503 (Apr. 2011).

361. Chan, C. T. Y., Lee, J. W., Cameron, D. E., Bashor, C. J. & Collins, J. J. 'Deadman' and 'Passcode' microbial kill switches for bacterial containment. *Nature Chemical Biology,* 1–7 (2015).

362. Chou, H. H. & Keasling, J. D. Programming adaptive control to evolve increased metabolite production. *Nature communications* **4,** 2595 (Jan. 2013).

363. Shis, D. L., Hussain, F., Meinhardt, S., Swint-Kruse, L. & Bennett, M. R. Modular, Multi-Input Transcriptional Logic Gating with Orthogonal LacI/GalR Family Chimeras. *ACS Synthetic Biology* **3,** 645–51 (2014).

364. Smillie, C. S. *et al.* Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* **480,** 241–244 (2011).

365. Wright, O., Delmans, M., Stan, G.-B. & Ellis, T. GeneGuard: a Modular Plasmid System Designed for Biosafety. *ACS synthetic biology,* 140513123533005 (2014).

366. Caliando, B. J. & Voigt, C. A. Targeted DNA degradation using a CRISPR device stably carried in the host genome. *Nature Communications* **6,** 6989 (2015).