THE UNIVERSITY OF CHICAGO

GENETIC ADAPTATION OF TWO ENTERIC BACTERIA TO SHIFTING ENVIRONMENTAL LANDSCAPES

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

COMMITTEE ON MICROBIOLOGY

BY

MANJING ZHANG

CHICAGO, ILLINOIS

MARCH 2022

To my wife, Lyndsee, whose love and support anchored me when I felt lost at sea.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction

All organisms require the ability to respond appropriately to changes in their environment. Inhabiting diverse environments (1) from dry deserts (2) to deep-sea hydrothermal vents (3) with only membrane layers separating their cytoplasm from the external world, bacteria are constantly interacting with their environments and being challenged by them to evolve. They adapt exceptionally quickly to these transitions via complex genetic, transcriptional, and post-transcriptional mechanisms (4–16). Studying the regulated and stochastic responses of bacteria to environmental pressures is significant to the advancement of our understanding of biological adaptation as well as the improvement of therapeutics for human health. In this dissertation, I present two studies investigating the ways in which enteric model organisms adapt genetically to host-associated selective pressures and discuss how the findings of these studies provide fundamental insights into microbial adaptation that can be used to improve clinical therapies for humans.

In the first story, we investigated extreme antagonistic gene-environment interactions in the enteric organism, *Escherichia coli*, using the clinically relevant antibiotic rifampicin. Extreme antagonistic gene-environment interactions occur between genetic and environmental perturbations that, individually, decrease fitness of the organism but, when presented together, increase the fitness of the organism. First, we addressed the lack of experimental methods to systematically identify such interactions by implementing a two-part strategy. Our strategy involves serial adaptive evolution by first adapting *E. coli* to a defined glucose medium, then adapting those strains to a sublethal concentration of rifampicin. Extreme antagonistic interactions address a fundamentally interesting question in biology concerning gene-

environment interactions, as well as provide targets for the design of antagonistic antibiotic combinations that can select against drug resistance (see: **Collateral Sensitivity**).

In the second story, we assess the shifting genetic requirements of a prevalent and abundant human gut microbe, *Bacteroides thetaiotaomicron* (*Bt*), during colonization of a germ-free murine gut. To understand how a bacterium ultimately succeeds or fails in adapting to a new environment, it is essential to know not only *which* genes are involved but *when* they are involved. The mammalian gut, which is regularly bombarded by exogenous microorganisms, represents a biologically and clinically relevant system to explore microbial adaptational processes. We introduce *Bt* into the guts of germ-free mice to 1) determine whether the genetic requirements for commensal colonization shift over time and, if so, 2) characterize the biological functions required for microbial survival at different points of colonization. The results of a high-throughput functional genetics assay (BarSeq), transcriptomics, and metabolomics converge on several conclusions. *First*, adaptation to the host gut occurs in distinct stages. *Second*, drastic shifts in gene usage occurred throughout the first week, shifting from high expression of amino acid biosynthesis to polysaccharide utilization genes. These changes were sustained thereafter, except for the continued upregulation of a single PUL responsible for the degradation of raffinose-family oligosaccharides (RFOs) rich in the standard chow diet fed to our mice. *Finally*, spontaneous mutations in wildtype Bt introduced to GF mice also evolved around this locus, highlighting the importance of efficient carbohydrate metabolism in long-term persistence within a monoassociated gut. To improve microbiome-based therapies, it will be important to appreciate and meet the distinct needs of the organism during the varying stages of colonization.

The Gut Microbiome

The human gastrointestinal tract (GI) harbors a dynamic and complex repository of microorganisms, which are collectively called the gut microbiome (17). The bacterial community that inhabits the human body varies across space and time (18). The intimate relationship between mammalian hosts and their resident microbes dates to the very beginnings of eukaryotic life on earth, and it has been posited that bacteria played an integral role to the dissemination of humans across the world by allowing our ancestors to survive in new ecological niches with diverse nutrient availability and pathogenic challenges (19). The gut microbiome is important for many aspects of host physiology, including resistance to pathogen invasion (20–23), immune modulation (20, 24–31), and metabolism (32–40).

The total genome of commensal microbes has been estimated to contain 150 times more genes than the human genome, equipping the host with essential functional traits that human beings have not evolved on their own(41). For instance, the carbohydrate-active enzymes (CAZymes) encoded in the microbial glycobiome allow the host to extract energy from otherwise inaccessible dietary fibers (35, 40, 42, 43). Further supporting human nutrition is the host of genes responsible for biosynthesis of vitamins, cofactors, and secondary metabolites (44, 45). However, not all host-associated bacteria are friend—a rare, but arguably better-studied, population represent pathogenic foe. These bacteria are often treated with oral antibiotics. Oral use of antibiotics leads to profound perturbations in the diversity of gut microbiomes and enrichment of antibiotic resistance genes (46–48). Emergence of antibiotic resistance is a major global health concern. At a concentration of 10^{10}—10^{11} bacteria g^{-1} of intestinal content, the gut microbiome represents one of the densest and most rapidly evolving ecosystems on Earth, which in turn, enables it to act as a highly effective reservoir for antibiotic resistance genes (49–52).

Therefore, given the important functions that the gut microbiome performs and the harm it causes when it is dysfunctional, the gut microbiome represents a relevant and clinically significant environment to study microbial adaptational processes.

Genetic Forces of Adaptation in Microbial Populations

Adaptation to selective forces depends on a myriad of variables, including but not limited to the intensity of environmental pressures, growth rate, mutation rate, and spatial partitioning. Though bacteria are many times more diverse and ancient than plants or animals, the evolution and adaptation of these populations has historically been much more poorly studied and understood than either that of plants or animals. However, this gap in knowledge is rapidly shrinking due to the technological advances of the past few decades and a new appreciation for the important roles that microbes play in the ecosystems on Earth as well as in human health. Below, I will highlight some considerations to frame our discussion of bacterial adaptation within a host-associated context.

Bacteria can make adaptations that are transient, semi-permanent, or divergent on an evolutionary scale (51–57). Most bacterial adaptations are what one would consider transient or semi-permanent, given the quick generation rate of most bacterial species and heterogeneity that is maintained even in populations under severe selective pressure. These factors allow rapid reversion of the adaptation when the stressor is lifted, whether through true genetic reversion or suppressor mutation (58, 60–62). The plasticity of gut microbes has been vital and strategic, historically, allowing changes in the lifestyle and dietary habits in the host, whether they be transient fluctuations or long-term cultural changes (43). Whereas the timescale of genetic

differentiation for plant populations is typically measured in years (63) and in large mammals in decades (64), natural bacterial populations and communities have been observed to develop genetic divergence in well under a year (65–67). In the case of replicate experimental populations, the timescale can be on the order of days (13, 15, 53, 68–71).

The plasticity of a bacterial genome is in part determined by its mutation rate. Mutation rate points to the raw occurrence of mutations and should be distinguished from fixation rate, which describes the speed by which the mutation pervades the population (72). Under weak selective pressure, most mutations are more deleterious than beneficial; thus, in theory, fixation rates should be minimized under these conditions (73). On the other hand, selection for beneficial mutations can be much stronger when bacterial populations face environmental transition. These populations may evolve a much higher fixation rate. In fact, clinical isolates of bacterial pathogens often have mutation rates 10- to 100-fold that of laboratory strains due to strategic defects in repair pathways, likely as a competitive evolutionary response against therapeutic agents like antibiotics (74, 75).

The development and increased accessibility of whole genome sequencing (WGS) have been instrumental to advancing studies of bacterial evolution. For example, *Mycobacterium tuberculosis* in its latent phase has long been assumed to have a low or non-existent mutation rate, since only 10% of patients with latent infection suffer reactivation, but when WGS was used to investigate this question in a primate model, latent phase mutation rate was found to be comparable to that of active *M. tuberculosis* (76). Furthermore, genomic analysis comparing *Pseudomonas aeruginosa* isolates during acute and chronic infection revealed that chronic isolates had a hypermutator phenotype which inactivate virulence genes and likely promoted long-term adaptation to the host (77). These findings advance our understanding of how

microbes may take similar or diverse approaches to adapting to different phases of colonization within the host and have serious implications for therapeutic treatment of these pathogens.

Moving beyond raw mutation rate, we should consider the factors that determine how quickly a beneficial mutation will pervade a population. In a well-mixed population, we would intuitively expect that that mutations will fix at a rate that is relative to the size of the growth-rate benefit. However, rarely do bacteria naturally live in a well-mixed environment (78–82). The niches that host-associated bacteria inhabit are almost never well-mixed, as they will form structures like biofilms on solid substrates like teeth (83) or inhabit distinct sections of the GI tract (84). Organisms living within a spatially structured habitat are generally thought to have slower rates of adaptation because competitive interactions are more likely to be localized (85–88). Thus, spatial partitioning becomes a key determinant to the fixation rate of new beneficial mutations.

Another factor key to determining the rate of evolutionary responses to environmental change is how quickly the environment is changing. At the same time that beneficial mutations arise and are being fixed in the environment, the environment itself may be changing on a gradient. Theoretical studies suggest that lower rates of change are associated with weaker selection (89, 90). Experimental evolution experiments in *Caenorhabditis elegans* demonstrate that under fast environmental change, genotypes with high fitness in the most extreme environment are quickly selected for; however, when the environmental change is slow and gradual, populations initially select for genotypes favoring the intermediate environmental condition and can lose the genotypes best suited for the most extreme environment due to genetic drift and founder effects (81). Corroborating this finding, a study on the evolution of antibiotic resistance in *P. aeruginosa* suggests that gradual increases in antibiotic concentration allow for more rapid evolution of antibiotic resistance but limit the extent of resistance, whereas steep increases in antibiotic

concentration slow development of antibiotic resistance but select for bacteria with resistance to higher doses of antibiotics (92). Together, these findings suggest that the rate of environmental change affects the pace and outcome of adaptive evolution.

Of course, microbes must strike a balance between genomic plasticity and stability (93). From one generation to the next, bacterial genomes are remarkably stable but are shaped by the effects of genome rearrangement, horizontal gene transfer (HGT), and the factors mentioned above that make them extremely plastic on an evolutionary scale. The evolutionary plasticity of bacterial genomes is a double-edged sword that thwarts our attempts to control pathogenesis as much as it allows commensal bacteria to shift flexibility to fit the lifestyle of the host. Providing stability to the bacterial genome is a host of robust repair, offense, and defensive apparatuses such as restrict-modification systems (94, 95). On the other hand, mobile elements, HGT, and genome rearrangement introduce variations to the genome and allow for the type of evolution in quantum leaps that bacteria are famous for (96). Genomic instability and environmental adaptation can occur through nearly countless other mechanisms, a few of which are discussed in the following section.

Molecular Mechanisms Supporting Environmental Adaptation

For all organisms, rapid and robust adaptive responses are essential for life in dynamic environments. Bacteria inhabit enormously diverse niches, and as such, have developed tight control of global regulatory circuits that coordinate metabolic flux with growth (97), cell envelope composition (98, 99), and cell cycle progression (100) to preserve viability in response to frequently changing conditions. Defects in these systems render bacteria unable to receive

cues and/or respond appropriately to them. This regulation can happen at any level of the central dogma of biology, from DNA to protein. Mutations involving deletions, duplications, and insertions change the amount of information contained in the genome. They can also result in the appearance of new junctions, which have the potential to alter the expression, and therefore function, of downstream genes. Point mutations can disrupt or alter the function of the encoded protein. Depending on the location of the disruption, mutations might cause loss of function, gain of function, or be silent (101). The effects of some mutations are local, whereas others can cause global reprogramming of the transcriptome (102–104). In many cases, biological functions emerge from interactions rather the sole contribution of a gene or environmental factor, as in synergistic or antagonistic gene-gene or gene-environment interactions (105–108).

Adaptive evolution often occurs in two phases: a "plastic" phase in which the environmental transition induces phenotypic changes without genetic mutation, and second phase in which accumulating mutations induce more permanent phenotypic changes (109). Plastic adaptations are particularly helpful in fluctuating environments, which describe most host-associated niches including the human gut. Bacteria respond rapidly and reversibly to new environments through a diverse arsenal of complex transcriptional and post-transcriptional regulatory mechanisms. Collectively, we refer to the changes in gene expression that allow bacteria to adapt to changing or suboptimal environments as stress responses (110). These responses are often regulated by a specialized sigma factor, two-component system, or some other transcriptional regulator (88–92).

The stringent response is a ubiquitous stress signaling pathway used by virtually all bacteria to survive nutrient starvation (116, 117). Under nutrient starvation, the guanosine penta- or tetra-phosphate alarmone commonly referred to as (p)ppGpp accumulates and reshapes the metabolism and physiology of bacterial cells. The widespread range of (p)ppGpp targets and

diversification of function allows bacteria from vastly different environments to respond to stress in ways that are relevant to their niche and lifestyle. However, some common consequences of the stringent response include a decrease in DNA replication and rRNA synthesis, and a concurrent increase in transcription of amino acid biosynthesis and nutrient transporters. The stringent response has been reported to mediate virulence gene expression and survival in pathogens (118–120). The human commensal *Bt* also relies on the stringent response to survive carbon starvation and persist in the gut through the stringent response (121).

Using WGS from bacterial isolates sampled at different points of clinical infections, a recent study reported on a common group of genes that undergo adaptation due to host association (122). Though it is often difficult to find direct orthologous relationships between specific genes undergoing adaptive changes, the authors found similar bacterial cellular processes evolving in a host-associated context. Interestingly, transcription factors were prominent among genes undergoing adaptive change. This suggests that transcriptional remodeling is one common mechanism to host adaptation. Another study followed up on one of the transcription factors in *Bt* and found that it is required for utilization of multiple polysaccharides as well as murine gut colonization (123). It will be interesting in future studies to determine whether these adaptive requirements result specifically from a lab-to-host transition or if even host-to-host transitions would also require global transcriptional reprogramming.

In the second phase of adaptive evolution, mutations take hold within bacterial genomes. These genomes are a mosaic of stable and unstable regions, crafted by genetic events like recombination, insertions, deletions, mutations, duplications, inversions, and transpositions (93). The most common type of mutation is point mutations, which are regarded as the raw material of evolution (124). Single-nucleotide substitutions in DNA may produce a synonymous codon

mutation or non-synonymous mutation. Point mutations in non-protein-coding DNA sequences can have significant functional consequences if they occur within regulatory regions of the genome. Recombination is another common mechanism through which new genetic variants are produced. It has dual roles in protecting genome stability as well as increasing genome flexibility by introducing new genetic information through incorporation of loci near identical or near-identical DNA sequences. Gene deletion also shaping bacterial evolution. While bacteria increase their DNA content through HGT and gene duplication, they are constantly purging non-functional sequences through deletions (125). Gene reduction by deletion can provide a selective advantage by eliminating functions that are dispensable and/or redundant, thus conserving an organism's limited resources (101, 102). This is not to say that insertion/duplication events are uncommon or insignificant to the fitness of bacteria. In fact, the occurrence of tandem duplications is prevalent among bacteria, and an increasing number of studies have demonstrated the examples of bacteria using tandem repeats to reversibly shut down or modulate the function of specific genes. For example, the phase-variable acetylation of the K1 antigen of *E. coli* K1 is driven by a heptanucleotide TR tract. Loss or gain of repeat units not in a multiple of three results in NeuO expression, which enhances desiccation resistance but reduces biofilm formation (128, 129). Tandem repeats also create variation that contributes to adaptation in the face of antibiotics for *S. haemolyticus* (130). This mechanism has also been noted in *S. enterica*, for which the amplification of the isoleucyl-tRNA synthetase gene facilitates resistance to mupirocin (131). However, due to the high intrinsic instability of these elements and the fitness cost of maintaining redundant copies, these genomic changes tend to be transient, which is why they have been underappreciated in the literature despite their immense and significant role in shaping microbial genomes.

Tandem repeats are often the result of transposable elements (TE) (132). TEs represent a special type of recombination in which DNA can migrate to a non-homologous locus. The type of transposable element that makes tandem repeats are replicative in that they cause DNA replication and produce a cointegrate, but they can also be conservative and only cause insertion without creating an additional copy of itself (133). While some transposons are site-specific, others are less stringent with their target sequence. Transposons can travel on plasmids and be transferred through HGT. (101)The contribution of regulatory elements like promoters is often overlooked in the literature. A recent study finds that the promoter activity of two *Drosophila* transposons in the Mariner superfamily can drive transcription of a reporter gene in a distantly related organism, including in bacteria (134). Orientation and proximity of the insertion to promoter sequences can have major consequences for the genomic impact of the IS, particularly if the insertion occurs upstream in an operon. Observation agrees with intuition in that these movable elements occur more frequency in bacteria residing within adverse conditions. Their contribution to altering pathogenic properties of infection including *S. epidermis* (135) and *M. tuberculosis* (136, 137) is well-noted in the literature.

Combinatorial Mutations

Some of the most interesting and complex adaptive mechanisms arise from interaction between multiple mutations. Indeed, recent advances in network modeling suggest that a single cellular function is rarely carried out by a single gene, and that this is a trend found not only in microbiology but generalizable across the life sciences (138–143). For our null hypothesis, we assume two non-interacting mutants, and define a "neutral" phenotype as one in which the

double mutant has a phenotype that is simply additive of the single mutants'. A double mutant with a more extreme phenotype than the neutral phenotype defines a synergistic, or synthetic, interaction. Alleviating interactions are defined by a double-mutant phenotype that is less severe than expected and may suggest that the gene products operate together or serially within a pathway. In the case when the first mutant already impairs the entire pathway, adding another mutant within the same pathway will have "diminishing returns" to the phenotype (144). These alleviating interactions can also be considered antagonistic, as the fitness impact of the second mutation is directly limited by the effect of the first.

Collateral Sensitivity

A unique case of antagonistic interactions is synthetic rescue, in which two genetic perturbations that are each individually deleterious become beneficial in the presence of each other (145). This can be considered an extreme form of antagonistic interactions because the double mutant not only confers *less negative* fitness but flips sign and confers *positive* fitness to the organism. These types of interactions are not constrained to gene-gene combinations, but can be between genetic and environmental perturbations, or two environmental perturbations. Antibiotic drugs are a clinically significant class of environmental perturbations. Given the current global crisis of antibiotic resistance, there is an exigent need for new treatment strategies to combat development of resistance. Among the most promising approaches is collateral sensitivity—a phenomenon in which acquisition of resistance to one antibiotic causes increase sensitivity to a second antibiotic (146). Despite previous efforts to identify extreme antagonistic drug pairs (147–151), there has been a lack of a systematic strategy to find these rare combinations. Finally, we need to address a

major knowledge gap of understanding the molecular mechanism of collateral sensitivity by identifying the transcriptional changes that represent the adaptive response and the cost of resistance.

Sources of selective pressure within the mammalian gut

Recent advances in sequencing technology have allowed us to appreciate the rapidly evolving microbial populations in the human gut. This finding is important for basic and translational gut microbiome research, but our understanding of the selective forces that act on these communities is still lacking. Below, I outline some biotic and abiotic pressures that shape the diversity and function of the microbiota.

Nutrient availability

Among these variables, the metabolic landscape within the gut has proven to be among the most critical and complex. A major nutrient source for resident bacteria is the plant-derived complex polysaccharides that are provided through the food consumed by the host. Since these fibers are indigestible by the host, they reach the gut lumen at higher concentrations than simple sugars and are thus more a more abundant nutrient source for the gut microbiota (152). Host-derived polysaccharides, such as mucus glycans, provide an alternate nutrient source for microbes that resident near the mucosa. Commensal *Bacteroides* species possess an impressive arsenal of polysaccharide utilization loci (PULs) to metabolize a wide variety of plant- and host-associated glycans as well as transport the generated simple sugars. These PULs are often integral in

determining the niche adaptation of *Bacteroides* species in the gut; as such, there is complex and fascinating regulation of these genes to facilitate quick adaptation to the environmental niche (38, 99).

Metabolic genes have been directly linked to commensal colonization. The commensal *Bacteroides fragilis* encodes for a unique class of PULs called commensal colonization factors (CCF), which are upregulated at the colonic surface and allow *B. frag*ilis to penetrate the bottom of colonic crypts. The CCF facilitates colonization through a combination of regulating the biosynthesis of capsular polysaccharide genes, whose expression is also intimately linked with nutrient availability; inducing PS-specific IgA responses that improve *B. fragilis* adherence to intestinal epithelial cells; and possible usage of a specific polysaccharide substrate by a putative chitobiase (153). Furthermore, hybrid selection RNA sequencing shows that *B. fragilis* colonized in the mucosal niche is significantly more metabolically active than *B. fragilis* localized to the lumen, upregulating a number of important genes for mucin glycan foraging (154). Similar observations have been made in *Bt*, linking CPS expression with polysaccharide utilization, host immune response, and gut colonization (38).

Not only do gut microbiota deploy a specific genetic program for metabolism at different biographical sites of the body, but it has also been shown in comparisons between long- and short-term dietary changes that timing is another variable that affects the way diet modulates bacterial adaptation. Candela *et al*. 2012 showed that within the gut microbiome, certain groups of bacteria were affected by short-term dietary changes while others were exclusively modulated by long-term dietary habits (155).

Antibiotics

Though only a small percentage of all bacteria are associated with human disease, those mighty few remain a matter of serious concern to global health. A few decades ago, during the golden era of drug discovery, there once existed the notion that all pathogens could be eradicated by chemotherapeutics and antibiotics. Currently, we live in an era of crisis in terms of the rapid emergence of resistant bacteria worldwide. The gut microbiome is considered a reservoir for antibiotic resistance genes from commensals to pathogens—called the "resistome" (156–159). This is in part in due to the density of this community, which facilitates HGT, as well as routine and liberal oral administration of antibiotics, which selects for organisms carrying antibiotic resistance genes.

Resistance has been developed to all current antibiotics, and with the pharmaceutical pipeline drying up in antibiotic development, there is a need to understand how we got here, and how to develop novel strategies to extend the lifetime of new drugs in the future. One of the greatest challenges that we face is the difference in the naturally rapid evolutionary rate of microbes (hours to days) compared to the laborious process of drug development (years) (159). In this way, it is clear to see that the spread of antibiotic resistance far outpaces that of new drug development. Given time, any organism living under selective pressure will inevitably evolve to resist the pressure and gain fitness within that environment, but there are ways of finding targets to impede the evolutionary process. One such approach is presented in **Chapter 2**.

Host immune system

The early life of a human presents a unique window of opportunity for educating the immune system. Favorable microbial exposures direct appropriate differentiation of immune cells, whereas missed opportunities within this window may result in inappropriate inflammatory responses to commensal microbiota later in life (24, 160).

Though it is assumed that proper development of the immune system provides for protection against pathogens and tolerance to commensal microbiota, the line between these groups can often in unclear (28, 30). The presence or absence of certain outer membrane components on commensal organisms can alter its relationship with the host in different ways. One well studied example of this is expression of specific CPS genes in *B. fragilis*. *B. fragilis*, much like many other *Bacteroides* species, can produce multiple CPS that comprise a microcapsule layer. One such CPS, PSA, has been the subject of many studies for its ability to stimulate an anti-inflammatory interleukin-10 (161–165). It has also been found in the gut microbe *Ruminococcus gnavus* that isolates with a protective isotype of its CPS promotes a tolerogenic immune response while isolates lacking function CPS genes promote inflammation within the host (166).

It is evident that an intimate and complex interplay occurs between microbes and the host immune system. Therefore, genes that promote or suppress host inflammation should be tightly regulated, as is the case for the well-studied CPS regulators in *Bacteroides*. Furthermore, when intestinal immune responses are dysregulated, they can result in chronic inflammatory disorders for the gut, such as IBD (167, 168) and food allergies (169, 170). These inflammatory environments can then, in turn, select for specific strains of microbes that are willing to withstand, or even thrive, in these conditions. A longitudinal study of ileal pouch patients

identified rapidly evolving genetic elements within *Bacteroides* species, including unique CPS loci (138). It remains unclear whether these adaptations represent attempts to stimulate, suppress, or evade host immunity—or if the evolution and selection of these loci are not a response to but a stimulus that incited inflammation.

Community effects

One of the most fascinating selective pressures within the gut is the fact that these bacteria live within a dense and diverse community of other microbes, with which they can develop cooperative, competitive, or neutral interactions. In many contexts, cooperative phenotypes have been observed to be central to the bacterial community functions such as quorum sensing (172), biofilm formation (173), and rapid spread of antibiotic resistance (158, 159). As mentioned previously, metabolite acquisition is a major driver of bacterial behavior and evolution. A fascinating example of cooperativity comes from a dedicated cross-feeding enzyme system between the prominent gut commensal *Bacteroides ovatus* and *Bacteroides vulgatus* (174). *B. ovatus* digests inulin extracellularly at a cost to itself and the benefit of *B. vulgatus*, and yet extracellular digestion of inulin increases the fitness of *B. ovatus* due to reciprocal fitness benefits it receives. Interestingly, despite cooperativity being a win-win for all members involved, natural cooperative interactions are rare and prone to instabilities on both ecological and evolutionary timescales as cooperativity renders cells dependent on species that may not always be nearby (166, 175–179).

On the other hand, ecological theory and modeling with recent data suggest that microbial competition can increase stability by dampening cooperative networks (175). Indeed, there are

many examples of competitive interactions among gut microbes, and bacterial antagonistic relationships are appreciated as a critical factor in shaping this complex system. One mechanism of interbacterial competition is the type VI secretion system, whereby effector proteins are injected into the target organism by the attacking cell (180–183). Competitive mechanisms such as these force the attacking and target species into an arms race, as in predator and prey.

However, the extent of bacterial adaptation that occurs due to selective pressures placed on individual species and pairwise interactions depend on the entire community context. A recent study demonstrated that community context can dramatically alter evolutionary dynamics in that there is a stronger evolutionary response in low-diversity communities, and that bacterial adaptation is, to some extent, constrained by the complexity of the surrounding community (184). In other words, while environmental factors provide the impetus for adaptation, individual species within that population only stand to capitalize on those opportunities if they have the capacity within a highly competitive (diverse) environment.

Bacterial Genetic Tools to Study Adaptation

Genetic Screens

Genetic tools are required to take full advantage of the wealth of information generated by genome sequencing efforts and ensuing global gene and protein expression analyses. Many elegant plasmid, cloning, expression, and mutagenesis systems have been developed over the years, and many are commercially available within model systems like *E. coli* and *Bacillus subtilis*. This is not true for most other bacterial species, which are still considered genetically

intractable due to many factors such as low conjugation frequency, severe barrier for foreign DNA entry due to its restriction-modification system, or lack of available selection markers within the system.

However, advances in sequencing technologies have made it feasible to capture WGS variation in organisms for which genetic manipulations are not yet available, such that they may be studied in their natural form. Other major advancements in the field of bacterial genetics include the increased variety and throughput of genetic screens. Genetic screens are critical tools for defining gene functions and assessing mutant organisms for phenotypes of interest. The process can be "forward" or "reverse". A forward genetics screen begins with the measurement of a phenotype followed by mapping of the causative genes. Reverse genetics starts with a known gene and assays for resultant phenotypes due to disruption of that locus. This approach can yield knowledge about the function of individual genes but is limited by hypotheses about the phenotypic outcome of the targeted mutant, and often involves assaying a single mutant in a wide variety of conditions, as in an OmniLog system (149).

Forward genetic screens generally involve generation of chromosomal mutants either randomly or site-specifically. Arguably the most powerful method for generation such a pool of mutants is through transposon insertion. Such an approach has been used to generate genome-wide mutant libraries in many model and non-model systems (186–190). A large challenge in microbiology is the sheer number of uncharacterized genes of unknown function that has been identified by genome sequencing. Transposon mutagenesis followed by high throughput screening of conditions is a powerful way to annotate gene function in bacteria, but many existing methods for TnSeq are often time consuming and require remapping of each library following the experiment. This process was linearized through the development of random barcode transposon-

site sequencing (RB-TnSeq) followed by barcode sequencing (BarSeq) within the experimental condition to monitor the abundance of mutants.

Experimental Evolution

Several studied within the past few years have capitalized on barcoding approaches in combination with experimental evolution to monitor bacterial adaptation occurring in real-time. For example, Lam and Monack (2014) infected mice with eight *Salmonella* strains with distinct DNA barcodes and observed that distinct subpopulations arose within the GI after 35 days, and early inhibitory priority effects excluded certain subpopulations from competing successfully for nutrients, thus being shed. Within this experiment, distinct sets of barcodes emerged from the individual mice, suggesting that stochastic forces or post-colonization drift occurred within these systems (191). Vasquez *et al*. 2021 took a similar approach, using DNA barcodes to track bacterial strain-level dynamics of *E. coli* in the mouse gut (68). In this set of experiments, mutations in motility and metabolic genes were reproducibly selected for. Furthermore, barcode tracking allowed them to use a population-genetics model to quantitatively predict transmission kinetics and determine that within-case migration accounted for a substantial fraction of the resident microbes in cohoused mice.

Metabolomics

The dynamic interplay between the gut microbiome and the host involves crosstalk between many different classes of small molecules (192). Together, these molecules contribute to the

evolutionary fitness of both microbe and host. While metagenomics and metatranscriptomics have been used extensively to characterize the gene and transcript contents of the gut microbiome, it is important to assess not only what genes microbes have and which transcripts they are making, but what metabolites are present within the GI milieu. Metabolomics has emerged as a technique that helps to characterize biological functions of host and gut, and most importantly, the impact of their intersection or dysregulation on risk for certain diseases.

Mass spectrometry-based metabolomics is a powerful way of detecting a broad range of small molecules within complex biological samples (193). This approach is supported by many sophisticated tools that have been designed to facilitate processing of the huge volumes of data that is generated (194). However, the fact remains that the bottleneck to the metabolomics studies often lies at the complexity of data processing. With the advancement of data processing tools, metabolomics will certainly become a more commonplace approach to facilitate exploration of the biological networks within the gut microbiome. Owing to the complexity of the microbiome, studies should increasingly aim to present multi-omics data as a way of cross validating the reported results.

Translational significance for studying adaptation of gut-associated bacteria

The gut is a fascinating environment for studying bacterial adaption owing to the dynamic interplay between host and microbe and among the complex community of microbes that together thrive within this dense and diverse system. Furthermore, the gut microbiome has great clinical significance, as it has been demonstrated to be a reservoir of antibiotic resistance genes, serve critical immunological and metabolic functions for the host, and be linked to a whole array

of diseases in dysbiosis. Probiotics and fecal microbiota transplantation are two major clinical interventions that involve colonization of exogenous bacteria to prevent or treat dysbiosis (195). However, the efficacy of both interventions has proven to be inconsistent, owing to a myriad of poorly understood variables including the host environment (196–206). Microbiota interventions will likely not be a "one size fits all." Thus, it will be vital to understand the genetic determinants of bacterial adaptation within the gut while adjusting for community complexity and structure, host disease state, genetics, age, and across various timescales such that these studies will inform improvement of microbiome-based therapeutics in the future.

Chapter 2: Extreme Antagonism Arising from Gene-Environment Interactions

2.1 Preface

The contents of this chapter were modified and adapted from its published form:

Wytock TP, **Zhang M**, Jinich A, Fiebig A, Crosson S, Motter AE., *Biophys J* (2020)

DOI: https://doi.org/10.1016/j.bpj.2020.09.038

Data Sets may be found online.

Contributions: Under S.C. and A.F.'s guidance, M.Z. designed and performed all wet-lab experiments, analyzed the growth data, and mapped the mutants. The T.P.W. performed all transcriptional analyses and led writing of the manuscript with input from all authors.

2.2 Introduction

Importance:

Mutations that are deleterious in the absence of an antibiotic can become beneficial in its presence, which is an example of an extreme antagonistic gene-environment interaction. Such antagonism is of biophysical significance because it reflects nonlocal interactions mediated by intracellular networks rather than direct physical or chemical interactions. We develop and apply a forward-evolution experimental approach to systematically identify these interactions using DNA and RNA sequencing. These analyses reveal differences in the expression of the translational machinery between antagonistic and nonantagonistic mutations. Our findings demonstrate how distinct single-basepair mutations within the same gene can have divergent phenotypic consequences, which give rise to antagonistic interactions that can be explored in addressing antibiotic resistance.

Background:

Complex biological functions generally emerge from interactions rather than solely from the contributions of individual genes and environmental factors (131). When the function of interest is growth rate, as often considered in fitness studies of single-cell organisms and immortalized cell lines, interactions are typically characterized in terms of whether they enhance or suppress fitness beyond what is predicted from the individual contributions. In the case of gene-gene and drug-drug interactions, pairwise relationships have been systematically identified using gene

knockouts and other function-impairing perturbations that effectively query combinatorial effects, through large-scale genetic screens (132–136), metabolic modeling (137–139), and collateral sensitivity assays (122, 139, 140). Interactions are classified as synergistic if the double-perturbation fitness is lower than expected from adding the effects of the individual perturbations and antagonistic if higher (144). Synergistic and antagonistic interactions are often interpreted as the result of local mechanisms (131, 141, 142). In this interpretation, synergistic interactions would relate genes in parallel or redundant pathways. Antagonistic interactions, on the other hand, would concern serial pathways of nonredundant genes or drug targets contributing to the same process. In the latter case, because the loss of either gene alone hampers the process, antagonism would emerge because the fitness impact of the second knockout is directly limited by the first.

An interesting exception to this local picture is the case of synthetic rescues, which are extreme antagonistic interactions in which a gene knockout becomes beneficial (rather than merely less deleterious) when applied after the knockout of another gene (143–146). Synthetic rescues can be interpreted as a form of sign epistasis (148) between two genetic perturbations that are each individually deleterious, in the sense that the fitness impact of one perturbation changes sign and becomes beneficial in the presence of the other. Although some synthetic rescues can be attributed to toxicity (149–154), previous work indicates that there exists a combinatorial number of synthetic rescues that are mediated by biochemical networks and involve genes in disparate processes (145, 155). Even though synthetic circuits have been used to study the evolution of extreme antagonism (156, 157), and the identification of gene-gene combinations exhibiting nonlocal, network-mediated synthetic rescues is ongoing (145, 146), an open proposition is whether analogous extreme antagonistic interactions can be systematically identified for gene-

environment combinations. Determining the prevalence of extreme antagonism would address a fundamental question in biology concerning the nature of possible interactions between genes and environmental factors. Furthermore, implementing such environmental factors using antibiotic stressors would offer a pathway to design new antibiotic combinations that exhibit collateral sensitivity, as discussed below, because the action of one antibiotic disfavors the acquisition of resistance to another (123, 124, 140, 157–161).

Here, we explore a method to systematically identify extreme antagonistic gene-environment interactions. These interactions occur between genetic and environmental perturbations that, individually, decrease fitness relative to the original condition but that, when combined, increase fitness relative to the environmental change alone. In parallel with synthetic rescues vis-à-vis sign epistasis, extreme antagonistic gene-environment interactions are an outstanding class of environmental sign epistasis (162), in the sense that mutations that are deleterious in a permissive environment become beneficial in a more restrictive environment. We identified these interactions using serial adaptive evolution of Escherichia coli, which consisted of 1) adaptation to a defined glucose medium yielding strains with enhanced fitness, and 2) subsequent adaptation of these strains to antibiotic stress defined by a sublethal concentration of rifampicin (rif). The gene-environment interactions were characterized by comparing the growth fitness of all strains grown in both the presence and the absence of rif, quantifying the cost of resistance (156). Previous studies of the cost of rif resistance have focused on strains cultivated in complex media and reported mutations in the β-subunit of RNA polymerase (rpoB) that result in a fitness disadvantage in such media in the absence of rif (164–166), though adaptive laboratory evolution experiments in defined glucose media have proved the same mutations to be beneficial (147). Because previously identified adaptive mutations that enhance growth rate and/or rif resistance

have pleiotropic effects (167–169), the global transcriptional consequences of adaptation were further investigated using RNA sequencing (RNA-Seq) before and after each adaptive step. Our experiments and analyses identified mutations in rif-adapted strains that confer growth faster, equal, and slower in the absence of rif compared to that of the parent strain. Of these three groups, the mutations conferring slower growth exhibit extreme antagonistic interactions with rif. These differences in growth can be attributed to widespread transcriptional reprogramming, which we map to broader cellular processes including central metabolism, translation, and a stringent-like response.

We propose that extreme antagonistic gene-environment pairs in which the environmental perturbation is the addition of a drug, such as rif in our experiments, provide targets for the design of antibiotic combinations that can select against resistance. For a pair of extreme antagonistically interacting antibiotics, growth inhibition by a drug combination is weaker than that of one of the drugs alone, meaning that acquisition of resistance to a second drug would cause growth to be suppressed more strongly (44, 45)—a phenomenon referred to as collateral sensitivity (140). Although previous studies of collateral sensitivity have found some extreme antagonistic drug pairs (122–126, 168), such pairs tend to be rare (160), and are found by directly testing existing antibiotics.

We apply our method to study antagonism as it relates to rif, for which there appear to be no antagonistic partner antibiotics (141). Resistance to rif, as with antibiotic resistance in general, induces pleiotropic effects on cell physiology that mitigate the action of the drug but are otherwise suboptimal (167, 169, 172), which we can characterize transcriptionally using RNA-Seq in the presence of the drug (to discern adaptative changes) and in the absence of the drug (to identify costs of resistance). We expect transcriptional profiles to be predictive of collateral

sensitivity, as has been shown for chemogenomic profiles (173), to the extent that both reflect the molecular mechanisms underlying adaptation and the cost of resistance. These efforts may offer another avenue to manage resistance in bacterial infections that are treated with rif, such as Mycoplasma tuberculosis (173, 174). Our approach to find extreme antagonistic gene-drug pairs is scalable and can facilitate the systematic design of antagonistically interacting drug combinations by providing targets—the gene mutations—against which a second drug may be developed. Ultimately, our results provide insights into fundamental and applied aspects of gene-environment interactions.

2.3 Results

Selection of mutations in multiple conditions and their growth consequences

To identify antagonistic gene-environment interactions, we undertook a two-step laboratory evolution approach in which we evolved the WT strain in a defined medium condition and subsequently introduced a sublethal antibiotic stressor (Fig. 1A, Fig. 2).

First, we serially passaged the WT strain in M9G, thereby selecting for mutants with enhanced growth rate in this defined condition. Two Ref strains were sequenced revealing a T1037P mutation in *rpoB*, which encodes the β-subunit of RNA polymerase (RNAP) or a C8Y mutation in *pykF*, which encodes pyruvate kinase (Fig. 1B). These mutations are known to enhance fitness in anaerobic conditions (175) and defined media (71), respectively. Then, each Ref strain was serially passaged in M9G + rif at a concentration that attenuates growth of both strains (Fig. 2). During serial passaging, rif-adapted mutants arose, increasing the overall growth rate of the bulk culture and restoring it to pre-rif levels. We isolated a set of 10 independent rif-adapted Mut strains, five from each Ref strain. The growth outcomes following each step of our lab evolution protocol are shown schematically in Fig. 1C and Fig. 2. In the context of our experiments, it is the relative growth rate in M9G between the Mut strains and the Ref strain that determines whether the genotype constitutes an extreme antagonistic pair.

The genetic changes in each Mut strain are reported in Table 1. Rif binds at the active site of RNAP and inhibits transcription by preventing translocation (176). We note that the mutations conferring enhanced growth rate in M9G found in the Ref strains—*pykF*(C8Y) and *rpoB*(T1037P)—are distinct from those conferring rif resistance in the Mut strains, even though one Ref mutation is in *rpoB*. Genetic changes that the Mut strains acquired during the

process of laboratory evolution in rif all occurred in *rpoB* and cluster at three previously described E. coli *rpoB* rif-resistance sites: (I) the RNAP active site where rif binds, (II) the fork domain which ensures base pair fidelity (177); and (N) the DNA entry channel (178). Mutations at sites I and II are reported to effect transcription slippage (177), those at site II can alter termination (179), whereas mutations at the N site decrease the open complex lifetime (178). On average, strains harboring mutations at site I have the fastest growth, followed by those at site II and site N. Because all mutations conferring rif adaptation occur in the same gene, we refer to these strains simply by their amino acid substitutions in rpoB from this point forward.

To evaluate whether the mutations associated with rif resistance result in fitness costs in the original cultivation condition (i.e., M9G without rif), we categorize the Mut strains by their growth relative to their Ref parents in M9G. The Mut strains may grow 1) faster, revealing that mutations are beneficial in both conditions, possibly reflecting incomplete adaptation of the Ref strains to M9G; 2) equally fast, revealing that the mutations are neutral in the original condition; or 3) slower, revealing that the acquired mutations are deleterious in the first condition, thereby indicating antagonism to rif treatment. These three groups are indicated respectively by orange, purple, and blue colors in the subsequent figures and tables, such as in Fig. 3, which presents the growth rates of each strain, and Table 2-5, which presents doubling times.

We identified rif-adaptive mutations in each of the growth categories in Fig. 3. The fast-growth mutations consist of the uncharacterized, but previously observed (180), S508P mutation [derived from the *rpoB*(T1037P) parent], and the well-described mutations (176, 177, 181, 182) at the RNAP active site, D516G and H526Y [derived from the *pykF*(C8Y) parent]. Compared to their Ref parent, these mutations boost growth rate by 8, 15, and 16%, respectively. Although the set of mutations we observed are not exhaustive, growth-enhancing active site rpoB mutations

were not observed in the *rpoB*(T1037P) Ref strain. The neutral growth mutations—L511R and I572N in *rpoB*(T1037P) and I572S in *pykF*(C8Y)—exhibit few transcriptional similarities, and the existing similarities mostly occur between the Mut strains deriving from the same parent (Table 6).

Correlations between transcriptional responses to the rif-adaptive mutations

To define the transcriptional changes associated with the mutation combinations in M9G with and without rif, we measured transcript levels in all Ref and Mut strains using RNA-Seq and determined which genes were differentially expressed between each Mut strain and its Ref parent in both M9G and M9G + rif. We calculated the $\log_2$ fold changes in each case and the Pearson correlations between the fold changes in all genes for each pair of conditions.

There are five major trends in the transcriptional correlations shown in Fig. 4. Specifically, 1) all Mut strains are more highly correlated in the presence of rif than in its absence; 2) the mutations at the I572 residue are all correlated, with I572F and I572N being strongly so; 3) the five fastest-growing Mut strains (relative to their Ref parents), S508P, D516G, H526Y, I572N, and I572S, are correlated in the presence and absence of rif; 4) the Q148L and L511R Mut strains derived from different Ref parents have strongly correlated transcriptomes; and 5) all the slow-growing Q148L and L511R Mut transcriptomes are anticorrelated with those of the fast-growing H526Y, D516G, and S508P strains in both conditions. Finding (1) suggests that all Mut strains are in similar biochemical/physiologic space when dealing with the challenge of rif, whereas finding (2) raises the question of how much the Ref strain mutation determines the transcriptional response versus the chemical similarity of the substituted amino acid. Together, findings (3), (4), and (5)

imply that mutants that grow faster than their parent Ref converge on a metabolic/physiologic profile that is distinct from mutants with slower growth. The correlation values associated with Fig. 4 are included in Table 6.

Patterns in the annotations of differentially expressed genes

To discover functional patterns at the level of gene regulation, we performed annotation analysis using the PANTHER annotations (184), COG annotations (185), modulons (185), and origons (186) from RegulonDB (187). Of the four annotation schemes we consider, the origon analysis is the most fine-grained because it links transcription factors directly with the regulated genes.

We visualize the origon results in Fig. 5, using a network representation. There are two types of network edges: those that connect Mut strains with their differentially expressed origons in M9G and rif and those that connect two origons. The latter edge type indicates that the genes downstream of the master regulator in the origon at the edge tail are contained within the larger origon at the edge head. When considering genes included in origons, we observe all strains except H526Y share *phoP* downregulation in rif, whereas *rpoB*(T1037P)-derived strains also share *iscR* upregulation. In the absence of rif, the cluster I fast-growth strains share *zur* downregulation and *acrR*, *torR*, and *slyA* upregulation, whereas the I572F and I572N strains share *acrR* and *argR* upregulation, and the Q148L strains downregulate *phoB*.

The origons that are differentially expressed are generally all implicated in *E. coli*'s stress response, a trend corroborated by our analysis of modulons. The membrane, GadEWX, and His-tRNA modulons, the latter two of which are implicated in acid stress response and histidine

metabolism, respectively, tend to be upregulated in the presence of rif in all strains. Simultaneously, iron utilization and CysB, which are implicated in inorganic sulfate utilization, are generally downregulated. Overall, the results in rif exhibit the convergence in transcription observed in Fig. 4, that is, the transcription factors that Mut strains express in M9G + rif are more similar than those expressed in M9G (see Table with details online). In the absence of rif, the modulon activation for slow-growth strains show upregulation of RpoS and GadEWX, which is generally anticorrelated with that of fast-growth strains. The anticorrelation is particularly strong in the *pykF*(C8Y) background, whereas the Q148L and I572F strains show few similarities with each other despite their similar growth phenotype.

As several of the most strongly differentially expressed modulons are in stress response genes, we analyze genes in the stringent response regulon in Fig. 6, given that genes in this regulon were implicated in the evolution of enhanced growth rate in previous experiments (223, 255). The *E. coli* stringent response downregulates genes involved in ribosome synthesis and upregulates genes for amino acid biosynthesis to account for poor nutrient conditions (188). For each transcriptional profile, we calculated a stringent response score by counting the number of genes that change in accordance with their behavior during the stringent response (see Materials and Methods). The fastest-growing Mut strains (D516G, H526Y, and S508P) have transcriptional profiles that resemble the stringent response in the absence of rif, whereas the slower growing strains appear to oppose, or not exhibit, this response. This pattern is particularly pronounced in the *pykF*(C8Y) background but is not apparent in the *rpoB*(T1037P) background.

As expected, strains that have a stringent-like transcriptional profile also have decreased transcripts corresponding to ribosomal proteins and 5S rRNA (Fig. 7). These strains show downregulation of genes classified in translation, nucleotide synthesis, and amino acid synthesis

COG groups, as well as exhibit *fnr* activation and significantly enhanced expression of glycolytic genes *eno*, *pgk*, *pykF*, and *tpiA*. Together, these trends indicate a shift toward overflow metabolism in the faster-growing strains, wherein the majority of ATP production is shifted from oxidative phosphorylation to glycolysis (266). Meanwhile, transcriptional changes in fast-growing strains are anticorrelated with those observed in the slow-growth mutants L511R and Q148L—including three of the four extreme antagonistic cases—possibly reflecting a maintenance of oxidative metabolism, which is more energy efficient (Fig. 6). Thus, slow-growth and fast-growth mutations appear to restore growth through different means, and transcriptional changes that emulate those observed in the stringent response are associated with fast growth in the *pykF*(C8Y) background.

We analyzed the rRNA and ribosomal protein transcript data for all strains (Fig. 7). Rif-adapted mutants harboring I572 mutations and those growing faster than their parent in the absence of rif exhibit decreased amounts of ribosomal protein transcripts in that condition. Increased tRNA and 5S rRNA expression may indicate increased readthrough from more upstream genes in the rRNA locus; decreases in expression at the rRNA locus may likewise provide evidence for transcription termination. Elevated expression of amino acid biosynthetic genes may reflect a short-lived open complex. We note that changes in rRNA (see counts table online) may reflect an imbalance in ribosomal components, analogous to suboptimal protein/DNA ratios observed in previous extreme antagonistic drug interactions generally (243) and in rif specifically (267).

The varying transcriptional patterns of genes encoding ribosomal components motivated us to more closely examine the reads that mapped to the rRNA loci. This effort was aimed at determining whether the mutant RNAPs exhibited differences in transcriptional efficiency along the operon. In Fig. 8A, we plot the fold change in read alignment averaged across all rRNA

operons for cells cultivated with and without rif. In the presence of rif, we see that both Ref strains have a pronounced peak in the 23S rRNA genes, 720 bp downstream of the rrl transcriptional start site. We note that the height of the peak correlates with the cost of the Ref mutations in rif as the *rpoB*(T1037P) mutant has slower growth in rif than the *pykF*(C8Y) mutant and a corresponding higher peak. On the antisense strand, we observe a peak immediately downstream of this site. Rif-resistance mutations appear to decrease the peak height (Fig. 8B), but the extent to which they do so is correlated with growth only in the *rpoB*(T1037P)-mutant-derived strains. When cultivated without rif, the change in number of aligning sequences at this site is positively correlated with growth (Fig. 8C). Although extreme antagonistic mutations appear to constitutively decrease the number of alignments at this site, the fast-growth mutations show substantially decreased alignments only in the presence of rif.

We compared the peak locations to those of pause sites (190) and rho-dependent termination sites (191). There is a rho-dependent termination site on the antisense strain 784 bp downstream of the rrl start on the sense strain and a pause site 832 bp downstream of the rrl start proximal to the 720 bp peak. In addition, a secondary site 1349 bp downstream from the rrl start recapitulates the behavior of the primary peak in response rif treatment. When comparing transcription of both Ref strains to their derived Mut strains in rif, decreases in alignments are apparent, but the ordering of strains is distinct from that observed at the 720 bp site. In the absence of rif, alignments increase by a larger amount in the *pykF*(C8Y) background at this site, which also has a predicted rho-dependent termination site. Beyond these two peaks, the alignment trends show decreases in transcriptional readthrough at the end of genes in fast-growth strains.

2.4 Discussion

Our identification of extreme antagonistic interactions reveals the role of molecular mechanisms fomenting gene-environment antagonism, including connections with transcriptional pausing and the stringent response. The application to an antibiotic environment reveals that the rif-adaptive evolution mutations alter the kinetic and thermodynamic properties of RNAP to counter the drug action but hamper growth in the drug's absence—consistent with the pattern of resistance mutations being pleiotropic (171, 172) and suboptimal (166, 168). Our study also demonstrates a proof-of-concept approach for the discovery of gene-environment interactions that is scalable to multiple genetic and environmental conditions and sufficiently sensitive to identify extreme antagonistic interactions. A central aspect of our approach is the two steps of adaptive evolution, as schematized in (Fig. 2). The first step, preadaptation to the initial environmental condition before applying the environmental perturbation, facilitates the identification of extreme antagonistic interactions (192, 193). The second step, adaptation to an antibiotic environment, reveals mutations involved in extreme antagonistic interactions that constitute candidate targets for the development of new antibiotic combinations.

Preadaptation to a defined medium reduces the likelihood that adaptive mutations to antibiotic stress would enhance fitness in the absence of an antibiotic, as predicted by Fisher's geometric model (194, 195). This is a step that was only recently explored in *Pseudomonas aeruginosa* (147), and our transcriptional analysis extends beyond this effort. We distinguish our preadapted strains, which are kept in exponential phase, from strains deriving from the long-term evolution experiment in *E. coli* B, in which cells transition daily between exponential and stationary phase (69, 192) and from unadapted strains in previous studies focused on nonextreme antagonistic interactions (160, 196, 197). The impact of preadaptation is illustrated in Fig. 2A-B,

by the intersecting Ref and WT growth lines, which show that M9G-adaptive mutations can sensitize strains to the addition of rif. This finding stands in contrast with studies investigating random mutations (14, 198), collateral drug resistance (197), and nutrient changes (199, 200), which suggest that adaptive mutations desensitize strains to other stresses. Indeed, direct selection of mutations conferring antibiotic resistance from unadapted WT strains tends to yield mutations that are beneficial in both environments (201–203), including the D516G and H526Y mutations in *E. coli* cultivated in M9G (146). We suggest that the difference originates from a higher cost of selection (193) across environments of a different nature, which, in the case of our experiment, implies that the same mutation generally cannot provide fast growth in both environments (with and without rif). The compatibility between these contrasting results can indeed be appreciated by recognizing that previous studies have mainly compared the response to environments of the same nature, such as different antibiotics (122, 168, 196) or carbon sources (198, 199). Indeed, antibiotic stress causes a targeted impairment of the cell's biochemical network that does not necessarily have an equivalent in adaptation to a nutrient condition. Our demonstration that adaptation to a nonantibiotic condition sensitizes the organism to antibiotic stress is corroborated by recent experiments in *Enterococcus faecalis* (204), and in accordance with observations of nonreciprocating collateral sensitivity between carbenicillin and gentamycin (147), we anticipate that the order of adaptive steps matters.

A possible mechanism explaining this tradeoff between generalization and specialization is a difference in transcriptional dynamics (180, 205, 206). Generally, the transcriptional changes observed between Ref and Mut strains in the presence of rif highlight the transcriptome-wide impact of adaptation, whereas the changes in the absence of rif reflect the cost of resistance. The alignment changes in Fig. 8 suggest that termination could play a role in rif resistance,

consistent with the observed changes in termination caused by mutations at the D516, H526, and I572 residues of *rpoB* (180) These findings may relate to rif resistance in *M. tuberculosis* (173), as recent work in that organism has implicated termination efficiency as the mechanism underlying the cost of resistance (174).

The experimental requirements to identify extreme antagonistic interactions are arguably more involved than those required to identify synergistic and merely antagonistic ones. Although we expect extreme antagonistic interactions to be large in number, they are also expected to be comparatively rare in the context of all interactions. This renders "reverse" genetic screening approaches, such as those based on libraries of gene knockouts (131–135), impractical because of the combinatorial explosion. Our "forward" genetic approach identifies point mutations that have pleiotropic effects on protein function in essential genes, which is a task that falls outside the scope of typical high-throughput approaches, and it is potentially massively parallelizable given advances in continuous culture and reductions in the cost of sequencing. Although high-throughput knockout experiments are not easily scalable in conjunction with preadaptation, future implementations of our approach could involve hybrid forward and reverse genetic approaches combined with mutagenesis techniques like TnSeq (135, 206) to expand the range of possible extreme antagonistic responses.

2.5 Materials and Methods

Strain cultivation

All strains used in this study were derived from wild-type (WT) *E. coli* strain K12 substrain BW25113 (207). The WT strain was suspended in media supplemented with 0.4% w/v glucose (M9G) in triplicate and allowed to grow at 37°C with shaking for 12 h, at which point cultures were diluted to an initial optical density at 600 nm (OD_{600}) of 0.01. Growth and dilution proceeded in this manner for 21–28 days (216–611 generations) to identify cultures with an enhanced growth rate. At the end of this period of evolution, cultures were plated, colonies were recovered, and WGS was performed, yielding Ref mutants with the mapped mutations *pykF*(C8Y) and *rpoB*(T1037P).

Selection for spontaneous mutants that restore fast growth in the presence of rif

Mutations that enabled fast growth of Ref strains in M9G with 7.5 μg/mL rif (M9G + rif)—an inhibitory concentration—were selected according to the following procedure: Three independent 5 mL cultures of each of these three strains were cultivated in 20 × 150 mm borosilicate tubes and serially passaged with a starting OD600 of 0.01 for 10 days, shaking at 37°C inclined at a 45° angle at 200 rpm in an Infors Shaker. As before, strains were diluted every 12 h to ensure the cells remained in exponential phase. Spontaneous acquisition of fast growth (i.e., rif-resistance) mutations is random based on published data (209). We observed an enhanced growth rate of the rif-containing cultures, as expected, within 4–5 days (8–10 passages; ~60–70 generations). Enhanced growth rate of the culture was typically evident after 4–5 days.

On the final day of serial passaging, these fast-growing cultures were streaked onto Luria-Bertani agar to isolate single clones. Growth rate measurements of clones confirmed that these "Mut" strains grew faster in M9G + rif than the Ref strains from which they were derived. The genetic identity of the Mut strains was verified by WGS.

Growth rate measurements

To measure growth rates, overnight cultures were first started from freshly grown colonies in 2 mL M9G shaken at 37°C overnight. These starter cultures were diluted in triplicates to OD600 = 0.01 in 2 mL of M9G either with or without 7.5 μg/mL rif in 13 × 100 mm borosilicate tubes and were allowed to grow at 37°C inclined at a 45° angle at 200 rpm in an Infors Shaker. The OD600 was measured approximately every 20 min for at least four independent cultures. Cultures were grown in the absence of light to prevent it from inactivating rif. To obtain growth rates, the log-linear region of the growth curves was fit to the exponential growth equation $OD600(t) = 2t/D$, where t is time and D is the doubling time, measured in minutes. The mean and standard deviation were calculated for the doubling time ratio of each derived mutant to its corresponding Ref parent. Differences in growth were assessed using a t-test on the ratio of Mut strain growth to Ref strain growth to control for possible differences in media preparation Table 1.

Mapping rif-adaptive mutations by WGS

We isolated genomic DNA from all E. coli mutant strains presented herein using a standard guanidinium thiocyanate extraction and isopropanol/ethanol precipitation. Briefly, the DNA was randomly sheared, and libraries were prepared for WGS using an Illumina HighSeq 4000 (50-bp single-end reads). The WGS data from each strain were assembled to the E. coli BW25113 WT genome template, and polymorphisms were identified using the breseq analysis pipeline (210). Each sequenced genome library yielded an average of 26 million reads, resulting in average depth of coverage greater than 250 times.

Measuring transcript levels by RNA-Seq

We isolated RNA for sequencing from both Ref strains and their rif-adapted Mut strains in triplicates, in both M9G and M9G + rif; cells were grown and harvested across three separate days. The growth protocol to prepare cultures for RNA isolation was as follows. All strains were cultivated in M9G at 37°C; starter cultures were diluted to OD600 = 0.01 with or without 7.5 μg/mL rif to an OD600 of ~0.18–0.22. At that point, RNA was extracted by pelleting cells for 30 s and rapidly resuspending the pellets in 1 mL TRIzol (Invitrogen, Life Technologies, Carlsbad, CA). RNA was extracted from the TRIzol suspension using the manufacturer's protocol. The extracted RNA was next treated with Turbo DNase (Ambion, Life Technologies, Carlsbad, CA) and further purified using an RNA purification kit (Qiagen, Hilden, Germany). Absence of DNA contamination was confirmed by PCR, where the lack of PCR product (~100 bp in length) relative to a DNA-containing positive control was interpreted as evidence of DNA removal. RNA-Seq libraries were prepared with an Illumina TruSeq stranded RNA kit

according to manufacturer's instructions. Sequencing (50-bp single-end read) was performed on an Illumina HiSeq 4000. Transcript levels were mapped to the E. coli BW25113 WT genome in CLC Genomics Workbench 11 (mismatch cost = 2; insertion cost = 3, deletion cost = 3, length fraction = 0.8, similarity fraction = 0.8). Sequencing reads for all strains, plus and minus rif treatment, have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus database (accession number GEO: GSE136977).

Annotation analysis of transcriptional changes

Determination of differential expression

We compared the transcript levels measured in Mut strains with their respective Ref parents when grown in M9G or M9G + rif. Differential expression was assessed using DESeq2 with the "apeglm" shrinkage estimator yielding log2 fold changes and single-gene p-values (211). These results were further interpreted in the context of Protein Analysis Through Evolutionary Relationships (PANTHER) pathway annotations, PANTHER protein classes, Clusters of Orthologous Groups (COG) identifiers, COG groups, modulons derived from independent component analysis of gene expression, and origons in the genetic network of the Regulon Database (RegulonDB).

PANTHER analysis

The log2 fold changes were input into PANTHER version 14.1 using the *E. coli* annotations for PANTHER pathways and PANTHER protein classes using a 5% false discovery rate (FDR) threshold for significance (184).

COG analysis

The COG annotation scheme groups genes into five classes, which are subdivided into 23 categories comprising 2161 identifiers based on protein amino acid sequence (185). We interrogated the transcriptional log2 fold changes for enrichment of COG categories and identifiers using a bootstrapping approach. For all *n* genes associated with a term, we calculated the rank sum and compared this with the rank sums derived from N = 20,000 randomly selected sets of n genes. We calculated the number *X* of randomly generated rank sums that were smaller than the observed rank sum among the distribution of randomly generated rank sums and determined a p-value using $1 - |1 - 2X/N|$. After determining the p-values, we selected terms for further consideration using the Benjamini-Hochberg FDR procedure with a 5% threshold (212).

RegulonDB transcription factor analysis

The network of 212 transcription factors and 1814 regulated genes was defined in the 'generegulation_tmp.txt' file downloaded from RegulonDB (188). Annotation enrichment proceeded as before, with the caveat that fold changes were multiplied by the valence of the transcription factor-gene interaction before calculating the rank sum. In addition, any transcription factors passing the 5% FDR threshold but having $\log_2$ fold changes in conflict with their downstream targets were excluded from consideration. We record all significant terms and their signs in table online for all Mut strains in both M9G and M9G + rif.

Quantifying the transcriptional signature of the stringent response

We reanalyzed the data contrasting the transcriptional outcomes of (stringent response-inducing) serine hydroxamate treatment between WT E. coli K12 MG1655 and its *relA*-deficient mutant,

which has its stringent response disabled (213). Genes that were differentially regulated between the two strains over the 30-min time-course were determined as follows. Linear regression was used to fit the parameters m and b in a = mt + b, where a is the gene expression vector and t is time. The slope parameters mWT and mΔrelA characterize the rate of transcriptional change in each gene. Then, the difference ($\langle \boldsymbol{m}WT\rangle - \langle \boldsymbol{m}\Delta relA\rangle$) characterizes the difference in gene regulation between strains having and lacking the stringent response. Let the standard error of $\langle \boldsymbol{m}WT\rangle$ and $\langle \boldsymbol{m}\Delta relA\rangle$ be $\langle\langle \boldsymbol{m}WT\rangle\rangle$ and $\langle\langle \boldsymbol{m}\Delta relA\rangle\rangle$, respectively. Then the quantity $(\langle \boldsymbol{m}WT\rangle - \langle \boldsymbol{m}\Delta relA\rangle)/(\langle\langle \boldsymbol{m}WT\rangle\rangle^2 + \langle\langle \boldsymbol{m}\Delta relA\rangle\rangle^2)^{1/2}$ is normally distributed about zero with unit variance, allowing application of a z-test. This test results in 183 genes that are differentially regulated (117 upregulated and 66 downregulated) between the two conditions (FDR = 5%).

The stringent response-regulated genes are quantified as a vector v indexed by the genes, g, with vg = 1 if gene g is upregulated, −1 if it is downregulated, and 0 otherwise. In our data, each transcriptional response to the adaptive evolution mutations is $\boldsymbol{\Delta} s = \log 2(\boldsymbol{a}s/\boldsymbol{a}Ref)$, where s is an index over Mut strains, a is the gene expression vector with the gene index suppressed, and the logarithm is applied elementwise. We define the stringent response score, S, for a given Mut strain to be $S = \boldsymbol{\Delta} s \cdot \boldsymbol{v}$. Statistical significance of the scores was assessed by bootstrapping, in which transcriptional responses were shuffled according to their average expression level across both conditions. The elements of Δs were shuffled (with s held fixed) and S', the randomized stringent response score, was calculated 10,000 times. Using $\langle S'\rangle$ and $\langle\langle S'\rangle\rangle$ to denote the mean and standard deviation of S', respectively, the quantity $(S - \langle S'\rangle)/\langle\langle S'\rangle\rangle$ is normally distributed about zero with unit variance. Applying the z-test yields the statistical significance of the score

against a null model that accounts for the total amount of transcriptional change across all genes in a given Mut strain.

Origon analysis

Origons are groups of genes reachable from a single master regulator in a transcriptional regulatory network, where master regulators are defined as transcription factors that have no regulatory inputs from other transcription factors (187). Using RegulonDB, we found 82 master regulators and calculated the degree to which they are turned on or off. Let G be a subgraph of the regulatory network reachable from master regulator O and let xw be the log2 fold changes for all genes w. Define the sign of each edge, $s(u,v)$, to be 1 if u is an activator of v, −1 if u is a repressor of v, and sign(xu) otherwise, and further define an indexing function $h(s(u,v))=1+(1-I(s(u,v)>0))$, where I() is the indicator function, which is 1 if the argument is true and 0 if false. We are now able to calculate ywz, the contributions to the activation of O for each node w, recursively. If $w=O$, the contribution is $yO1=1$ and $yO2=0$ because expression of the regulator should contribute to evidence of its activation. For $w\neq O$, let $P(w)$ be the set of direct predecessor nodes of w, and define $|P(w)|$ to be the number of predecessors.

Then, $ywh(s(u,v))=\exp((1/|P(w)|)\sum u\in P(w)\log(ryuh(s(u,v))))$, where $0<r\leq 1$ is a diffusion parameter to attenuate the contribution of longer paths. We calculate y starting at O, proceeding to its immediate descendants and so on through the network. At each stage, we ensure that all $u\in P(w)$ have been calculated before calculating ywz. In the case of cycles, there exists a $u\in P(w)$ such that $u\in D(w)$, where $D(w)$ is the set of descendants (direct and indirect) of w. In

this case, we ignore the contribution from all edges (v,z), where z is the ancestor of u and a descendant of w, but v is only a descendant of w. The averaged contributions to the origon expression is $\overline{y_w} = y_{w1} - y_{w2}$, and the total origon expression is $\bar{y} \cdot \boldsymbol{x}$.

To compare the origon expression against the random expectation, we reshuffle the nodes in the transcriptional regulatory network according to their in-degree and out-degree. We grouped by each value of in-degree from 0 to 12 and binned nodes with in-degrees higher than 12 together (for a total of 14 bins), and we grouped out-degrees into 24 logarithmically spaced groups. Of the possible 336 bins, 82 are occupied. To randomize, expression values were randomly permuted between nodes in each bin, thus preserving the degree distributions under randomization. The results of the origon analysis are reported online.

Modulon analysis

Modulons are groups of genes that contribute to a particular cell function. We projected our transcriptional log2 fold changes onto gene loadings for 92 modulons (186) and compared the resulting projections with those generated by reshuffling the transcriptional data. Modulons were considered significantly differentially expressed if they passed a 1% FDR threshold. Differentially expressed modulons are reported online.

RNA alignments of ribosomal RNA genes

To examine the RNA-Seq data for signs of transcriptional termination and pausing, we used the wiggle track formatted (WIG) files generated by Rockhopper and the gene locations from NC_000913.3 to calculate a rolling sum (50-bp window) of sequence alignments throughout the genome for both strands. These were averaged over three replicates to generate an alignment profile for each experimental condition. Next, we averaged the alignments for each ribosomal RNA (rRNA) operon over three segments: from the start of the *rrs* gene to the start of the transfer RNA (tRNA) genes, from the start of the tRNA genes to the start of the rrl gene, and from the start of the rrl gene to the end of the *rrf* gene. Because the tRNA genes are of different lengths in each operon (from 335 to 447 bp), we rescaled the counts by 447/L, where L is the operon length of the tRNA genes, and linearly interpolated the counts at each base before averaging. Once averaged, these three segments were concatenated into a single vector and annotated with pause sites (190) and termination sites (188, 190, 191).

Chapter 3: Dynamic Genetic Requirements for *B. thetaiotaomicron* Gut Colonization in a Murine Model

3.1 Preface

Authors: Manjing Zhang*, Megan Kennedy*, Orlando DeLeon, Jacie Bissell, Florian Trigodet, Karen Lolans, Sara Temelkova, Katherine Carroll, Aretha Fiebig, Adam Deutschbauer, Joy Bergelson, Eugene B. Chang

*co-first authors

Author Contributions: M.Z. and M.K. designed the research and performed all mouse studies collaboratively with help from K.C. M.Z. performed the initial transcriptomics analysis, and O.D. completed the GSEA. M.K. performed the BarSeq analysis. M.Z. analyzed the metabolomics datasets, designed, and performed the growth assays on isolated mutants with help from J.B. and S.T., and performed mutational mapping of RB-TnSeq isolates and spontaneous mutants. Long-read sequencing was performed by K.L. and analyzed by F.T. M.K. led the writing of the introduction and M.Z. led the writing of the remaining sections of the paper, with contributions from M.K., O.D. A.F., J.B., and E.C.

3.2 Introduction

Importance:

Microbes regularly disperse across and adapt to new environments and ecological niches. A clinically significant microbial niche home to trillions of microbes is the mammalian gut. Temporal processes of microbial adaptation over the course of gut colonization are poorly understood on a genetic, transcriptional, and metabolite level. In this study, we leverage a three-pronged approach to characterize gut colonization as a dynamic process with shifting genetic determinants of microbial fitness. This study sheds light on host colonization by an organism that is prevalent and dominant across healthy human microbiomes, *Bacteroides thetaiotaomicron*, and not only identifies key pathways involved in colonization, but determines the timing of *when* these pathways are most vital to colonization success. The findings and approaches of this study can inform future investigations using various bacteria and experimental models to improve our understanding of the longitudinal process of microbial adaptation within unique environments such as the gut.

Background:

Fast adaptation is paramount to the survival of any species undergoing an environmental transition. For microbial taxa, which are frequently and rapidly dispersed across dramatically different habitats and microenvironments, processes of local adaptation may arise as primary determinants of microbial colonization success and resulting biogeography (214). The

mammalian gut is an environment regularly bombarded with a diverse array of exogenous microorganisms. As such, it represents a biologically and clinically relevant system to explore microbial adaptational processes.

Canonically, adaptation occurs by way of genetic evolutionary mechanisms: genetic variants have differential fitness in the new environment, which establishes variation in survival and reproduction, and ultimately leads to shifts in gene frequencies over time. Over shorter timescales, organisms may adapt by altering plastic gene expression profiles to optimize functional characteristics including resource use, growth strategies, and resistance to environmental stressors (215–217). Both processes fundamentally reflect changes in fitness over time. To understand how a microbe ultimately succeeds or fails in adapting to a new environment, it is therefore essential to assess the temporal dynamics of its fitness over the course of colonization.

In the context of host-microbe interactions, adaptational processes play out within a dynamic host environment, in which feedbacks between microbe and host can have significant consequences for the fitness and survival of both (218). Understanding genetic and transcriptional changes of exogenous microbes upon entering a novel host system is a critical step toward elucidating the shifting landscapes created because of this feedback circuit. While previous work has evaluated the genetic requirements for colonization success across a variety of microbes in the guts of both conventionally raised and germ-free mice, these studies have generally assessed fitness at only a single timepoint, and therefore yield an incomplete picture of the temporal processes of adaptation and colonization (219–221). Moreover, although host transcriptional responses to microbial colonization are regularly documented, few studies have

comprehensively assessed microbial gene expression profiles over the course of colonization (222–224).

In the following experiments, we introduce a human-derived commensal organism, *Bacteroides thetaiotaomicron* (Bt), into the guts of germ-free mice to 1) determine whether the genetic requirements for colonization shift over time and, if so, 2) characterize the biological functions required for microbial survival at different points of colonization. To identify the genes important for fitness in the context of gut colonization, we combine two complementary unbiased approaches: transcriptomics (RNA-seq), which reveals global gene usage patterns, and a functional genetics approach [BarSeq (221, 225)] to assess fitness consequences of gene disruptions at a global scale over the course of colonization. Finally, we evaluate spontaneous evolution of wild-type Bt in the gut to survey natural population-level fitness dynamics. Our results indicate that adaptation to the host gut occurs in distinct stages. During the earliest stage of colonization, genes involved in amino acid and vitamin biosynthesis are upregulated and, in some cases, play essential roles in survival of Bt. By the end of the first week, expression of these genes is downregulated, and expression of carbohydrate metabolism genes peaks. These levels are sustained for the rest of the two-week experimental period, except for the continued upregulation of a single PUL responsible for the degradation of raffinose-family oligosaccharides (RFOs) rich in the standard chow diet fed to our mice. Spontaneous mutations in WT Bt also evolve around this locus, highlighting the importance of efficient carbohydrate metabolism in long-term persistence within a monoassociated gut.

These experiments lay the groundwork for future delineation of shifting colonization pressures in various host backgrounds and microbiome compositions. We expect that these insights into the temporally dynamic stresses that microbes must overcome to colonize and persist in the gut will

prove invaluable to our understanding of microbial adaptation and the development of microbiome-based therapies.

3.3 Results

Both transcriptional and genetic fitness determinants shift over the course of Bt colonization and persistence.

To evaluate global transcription during colonization, we introduced wildtype (WT) Bt into germ-free (GF) C57Bl/6 mice and collected cecal contents at Days 1, 7, and 14 after colonization (Fig. 9A). After rRNA and host RNA depletion, the bacterial RNA samples were sequenced and compared in a pairwise fashion across D1-D7 and D7-D14. In parallel, to assess functional genetic requirements during colonization, we introduced a rich library of randomly barcoded Tn insertion (RB-Tn) mutants of Bt into four different cohorts of germ-free C57Bl/6 mice and collected daily fecal samples (Fig. 9A). Amplification and sequencing of the transposon barcodes reveals the relative abundance of each mutant in the library at each timepoint.

First, we addressed the question of whether there are differences in Bt gene fitness at different times after introduction into the mouse gut. We performed Principal Coordinates Analysis (PCoA) using Bray-Curtis dissimilarity on the relative abundance of the RB-Tn mutant strains within each mouse across the days of the functional genetics experiment and found that across multiple independent cohorts of this experiment, the mutant pool composition shifted across time in a deterministic pattern (Fig. 9B). PERMANOVA analysis showed that the data cluster significantly by experimental day ($p = 1e\text{-}4$, $R2 = 0.457$). Qualitatively, the largest shifts in PC1 occurred between D1 and D7, whereas the mutant pool changed less dramatically between D7 and D14. Consistent with the global changes in mutant abundance with time, clustering of WT Bt gene expression data from our transcriptomics experiment reveals that the largest changes in

expression occur during the first week after introduction of B. theta (Fig. 9C). The Bt expression profile from D1 clustered distinctly from that of D7 or D14 (p-value < 0.05), whereas D7 and D14 profiles were statistically indistinguishable. Together, these data suggest that a large shift occurs early during the first week of colonization such that different sets of genes mediate colonization and growth of the organism before and after the switch.

Amino acid, vitamin, and capsular polysaccharide biosynthesis are transcriptionally upregulated and functionally significant during early invasion of the gut.

To determine the scale of the observed shifts in gene usage, we compared the transcriptomes of WT Bt isolated on D1 and D7. Differentially expressed genes (DEGs) were identified for each pairwise comparison using the parameters logFDR < -3, |log2FC| > 2, and max group mean > 50 TPM. Even using these stringent criteria, we identified a staggering number of DEGs. 154 genes were significantly enriched on D1 compared to D7, while 359 genes were significantly enriched on D7 compared to D1 (Fig. 10A).

Next, we asked whether the genes with higher relative expression on D1 correspond to specific functions. Though the period immediately following invasion of an exogenous organism has been studied in pathogenic infection contexts, very early time points in commensal colonization have been neglected in existing studies. This is a critical time for the invading organism, in which it must quickly adapt to a different set of stressors unique to the host environment, including differences in nutrient availability, pH, salinity, and assaults by the host immune system.

Previous attempts to characterize Bt colonization within the gut suggest that remodeling of the outer membrane is one response to the broad array of new selective pressures in the extracellular environment as the cell enters a new host (226–228). As is typical of other Bacteroides, Bt dedicates a substantial portion of its genome to its 8 distinct capsular polysaccharide (CPS) biosynthesis loci, which exhibit complex gene regulatory mechanisms. It has been shown previously that Bt can dynamically change its CPS expression profile in vivo (226), and that CPS4 plays a critical role in colonization (227, 228). Indeed, we saw major differences in expression of CPS biosynthesis genes between D1 and D7 (Fig. S11A). CPS1, 3, 4, and 8 loci have higher relative expression during the first day following introduction of Bt, whereas CPS5 and 6 are significantly enriched at D7 and beyond (Fig. S11B). Interestingly, mutants with disruptions in CPS1, 3, and 8 loci were not significantly depleted during the early days of the functional genetics assay. Only CPS4 mutants showed severe fitness defects (Fig. S12). This could indicate that CPS1, 3, and 8 serve redundant functions, or that fitness corresponds to protein levels or activity that are regulated independently of mRNA levels. Nevertheless, these results suggest that regulation of CPS gene expression is a response to the environmental changes that Bt encounters upon initial invasion of the murine gut.

To gain a more comprehensive understanding of the gene pathways expressed during the acute phase of adaptation to the host gut, we mapped the genes in the Bt genome to the 181 metabolic pathways within the Kyoto Encyclopedia of Genes and Genomes (KEGG) catalog and performed Gene Set Enrichment Analysis (GSEA) for each pathway on the differences in transcript abundances between D1 and D7 (Fig. 10B). The global expression pattern of Bt during D1 is reminiscent of the stringent response, in which growth is inhibited under conditions of nutrient limitation in favor of amino acid biosynthesis.

Our analysis of gene expression changes revealed that pathways corresponding to the biosynthesis of many essential amino acids were enriched specifically at the D1 timepoint, including those for histidine, lysine, leucine, and isoleucine This aligns with the findings of Watson et al. (229), which identifies biosynthesis of essential amino acids as one of the main drivers of microbial colonization outcomes after fecal microbial transplant. Concurrently enriched at D1 were genes involved in the biosynthesis of biotin, which is a cofactor required for many reactions, including amino acid biosynthesis. These results are supported by our RB-TnSeq analysis: when we mapped the significantly depleted mutants (t-statistic < -3σ) on the first day of the functional genetics assay using KEGG Orthology (KO), we identified amino acid metabolism as the KO group with the largest number of significantly depleted mutants (Fig. 10C). These genes are identified by gold stars on the amino acid biosynthesis pathway map in Fig. 13A, which gives a clear visualization of the individual reactions whose gene expression is enriched at D1 compared to D7. We can see that biosynthesis of most amino acids have either multiple reactions in the pathway that are transcriptionally enriched at D1, or one step that is functionally essential. Not only do both our transcriptomics analysis and functional genetics screen support a key role for amino acid biosynthesis early in colonization, but metabolomic analysis of the cecal contents corroborates this finding as well. When we measured the levels of specific amino acids in the murine cecum before, post-D1, and post-D7 colonization, we found that amino acid levels were generally much higher D1 after colonization than D7 or D14 (Fig. 13B). This difference was especially profound and statistically significant for amino acids in the glycine-serine-threonine pathway, as well as aspartate. On the other hand, glutamate and proline accumulate within the cecum during the first week of the time course before being depleted in the second

week. The biosynthesis of glycine, glutamate, and aspartate was found to be functionally significant to the survival of Bt in the acute phase of colonization (Fig. 13A).

Interestingly, while amino acid biosynthesis was the predominant functional hit among significantly depleted mutants at D1 in the functional genetics assay, over the next two days, amino acid biosynthesis becomes less essential to cell fitness as Bt begins to resume its carbohydrate metabolism activities, around D3, after overcoming an initial period of starvation (Fig. 10C, Fig. 14).

A shift toward enhanced expression of diverse sugar metabolism genes occurs during the first week of gut colonization

Members of the genus Bacteroides are well-known for their ability to digest a wide variety of polysaccharides. According to the CAZy database, Bt possesses 359 glycoside hydrolases, 87 glycosyl transferases, 15 polysaccharide lyases, and 19 carbohydrate esterases. Among the DEGs enriched on D7 compared to D1, carbohydrate metabolism stood out as the most prominently enriched functional category (Fig. 15). After the acute phase of invasion, Bt shifts toward expressing a remarkable number of polysaccharide utilization loci (PULs). Of the 42 total PULs that were significantly enriched at any point of the experiment, 34 were enriched on Day 7 or 14. These upregulated PULs encode genes belonging to a variety of pathways, including glycolysis, galactose degradation, pectin degradation, and D-galacturonate degradation (Fig. 16). Included among those are also PULs known to be involved in degradation of O-glycans (15), which suggests that B. theta is foraging for sugars through digestion of the mucus layer. These results, along with a 10-fold increase in CFU in the latter half of the experiment (Fig. 17), suggest that

within a week after introduction into GF mice, Bt recenters its transcriptome from a stringent response-like expression profile to a response centered predominantly on metabolism of the available sugars within the host gut.

Upregulation of α-galactosidase activity confers a significant growth advantage to Bt in GF mice fed a standard RFO-rich diet

In stark contrast to the large transcriptional shifts that occurred in the first week after introduction, the transcriptional profile at D7 is mostly sustained through the rest of the experiment (Fig. 18A). In other words, the transcriptome of host associated Bt has largely stabilized by the end of the first week. The one salient exception in the second week is the continued upregulation of PUL24, suggesting that PUL24 may enable Bt to gain a competitive growth advantage as it proceeds to the persistence stage of engraftment. PUL24 has an operon structure typical of a PUL, with regulatory elements, transporter proteins and carbohydrate-active enzymes (Fig. 18B). Interestingly, the last gene of PUL24 encodes an α-galactosidase (BT1871), which is predicted to confer the ability to hydrolyze the α-1,6 glycosidic linkage in raffinose family oligosaccharides (RFOs), a major component of the fiber-rich diet that our mice were fed (231). Though this study focuses on understanding the shifting genetic determinants of colonization over time, we recognize that the effect of diet is intrinsic to the results of any gut microbiome study.

Our functional genetics screen confirmed the importance of PUL24: within the first five days after introduction of the RB-Tn mutant library into GF mice, the diversity of the community quickly collapses (Fig. 18C) due to strong positive selection for a small pool of hyperfit mutants.

RB-Tn mutants with insertions in PUL24 were among the most enriched mutants at the end of the two-week experiment, even after controlling for their relatively high abundances in the inoculum (Fig. 18D). Furthermore, most of the significantly enriched mutants had TN insertions in one of three operons: PUL24 (BT1871-1877), PUL39 (BT2851-2860), or BT3130-3134, all of which encode at least one α-galactosidase, situated at the tail end of the operon (Fig 18D). At the end of the Tn selection experiments, the populations were overtaken by mutants carrying Tn insertions upstream of these α-galactosidase genes (Fig 18E). This was true for all four independent experimental runs, which were performed months apart from one another (Fig 18F). To note is the fact that in vitro passaging of this library within BHI-S medium also resulted in significant, though not as dramatic, decrease in Shannon diversity over two weeks, but the mutant library was able to maintain its diversity when passaged in the presence of a specific-pathogen-free community (Fig. 19). Introduction of this mutant library in a GF mouse is frustrated by severe bottlenecking, but the rare population that penetrate the gut is stably maintained throughout the course of the next two weeks (Fig. 20. This difference suggests that gene-environment interactions for Bt is affected by the presence or absence of other community members.

This RB-Tn pool has previously been assayed in over 300 different conditions including distinct carbon or nitrogen sources or specific stress conditions (16). The mutants that we identified in as being hyperfit in GF mice exhibit a phenotype significantly deviant from WT only in two conditions: within GF mice and in defined culture with melibiose—a disaccharide of glucose and galactose—as the sole carbon source. In both these cases, these mutants exhibit a growth advantage. Indeed, when we isolated the most abundant strains from D14 of the RB-Tn experiment, we found that their growth rate was significantly faster than WT when melibiose is

the sole carbon source (Fig. 21A). In this condition, Bt must hydrolyze the alpha 1,6-bond to harvest and metabolize the monosaccharide sugars. The competitive advantage of these mutants cannot be attributed to either of the monosaccharides, as the mutant growth rates with these carbon substrates are indistinguishable from that of WT (Fig. 22).

Given that release of the monosaccharide sugars of melibiose depends on hydrolysis of an α glycosidic bond, we wondered if the fitness phenotypes in both mice and in vitro depended on overexpression of the α-galactosidase gene downstream of the TN insertion site. The original publication characterizing this Bt RB-Tn library (16) showed that polar effects can be expected in which readthrough from the strong promoter driving antibiotic resistance in the transposon can overexpress genes downstream from the insertion. To test the hypothesis that the transposon insertion enhanced expression of downstream genes, we grew WT and hyperfit mutants in melibiose medium and measured the expression of α-galactosidase genes. For this experiment, we used two mutants, one carrying an insertion in PUL24 and another carrying an insertion in the BT3130-3134 operon. We found that, in both cases, expression of the α-galactosidase downstream of the insertion but within the same operon was overexpressed by at least 10 times compared to WT (Fig. 21B). Meanwhile, expression of an α-galactosidase (BT2851) located outside the operons carrying the insertion was similar between WT and the mutants. We then overexpressed BT1871 from a strong, constitutively active promoter (P_{rpoD}) in WT Bt, and observed a three-fold increase in log-phase growth rate when melibiose was the sole carbon source (Fig. 21A).

Metabolomic measurements of a carbohydrate panel confirms that high concentrations of raffinose and its constituent sugars are among the most abundant substrates within the ceca of GF mice fed a standard chow (Fig. 21C). Raffinose and melibiose build up in the lower GI of

these mice at high concentrations until Bt is introduced. After 7 days, during which time Bt initiates overexpression of alpha-galactosidase genes, these sugars are depleted (Fig. 21C, Fig. 23A). In contrast, the other disaccharide product of raffinose, sucrose, which does not have a 1,6-alpha glycosidic bond, is consumed by the host and no significant changes are observed in sucrose concentration following the Bt colonization (Fig. 23B).

Together, we find that when mice are fed a standard high-fiber chow, the high concentration of RFOs accumulating in the lower GI tract creates an environment that strongly selects for Bt strains that can make efficient use of these sugars through increased expression of α-galactosidases. We infer that efficient carbohydrate metabolism—particularly for abundant dietary fibers—is a major determinant of population-level selective dynamics during the persistence phase of Bt engraftment within the gut.

Strong selection for efficient RFO metabolism leads to emergence of spontaneous Bt mutants with duplicates the BT1871 locus through a IS3-family transposable element

The transposon mutants in our functional genetics experiment were able to gain ~10-fold increase in α-galactosidase expression due to readthrough from an extremely strong, synthetic promoter. We wondered if the selective pressure for elevated α-galactosidase activity to utilize the α 1,6-linked sugars abundant in the diet would drive evolution of a spontaneous mutant with enhanced α-galactosidase activity. We inoculated three GF mice with WT Bt. After one week, they were each separated into individual cages. At the end of the six-week incubation period, we performed shotgun metagenomic sequencing on fecal material from all three mice. When we assembled the short-read sequences, we note that there was at least 2x coverage of the BT1871

for all the samples (Fig. 24A). The dip in coverage mapped to an IS3-family transposase, of which there are six other identical copies in the genome; hence, the corresponding coverage was diluted among the different copies. BreSeq analysis (210) identified three new junctions in the Bt genome indicative of genome rearrangements (Fig. 24B). All three junctions were formed between either BT1872 or BT1873 and a locus downstream of BT1871, which encodes for rRNAs, tRNAs, and a ribosome recycling factor. All three new junctions occupied a significant portion of the sequencing reads (60-73%), suggesting that the mutants carrying these junctions were sufficiently competitive in the gut environment to comprise a significant portion of the fecal Bt population.

IS3-family transposases are known to act by a copy-and-paste mechanism (233). We wondered if this transposase could have created multiple copies of the BT1871 locus. We isolated individual mutants from the bulk fecal material, cultured the isolates in complex broth (BHI-S), and performed long-read sequencing with MinION. In one of the isolates, we identified a tandem repeat of the BT1871 locus (Fig. 24C), which places the original copy of BT1871 downstream of a strong rRNA promoter. Like the RB-Tn mutants, the mutants isolated in this experiment all grew about five times faster than WT in defined minimal media with melibiose (Fig. 25). Interestingly, the six identical copies of this IS3-family transposase are often found adjacent to PULs in the WT Bt genome. This suggests that these transposable elements confer a mechanism to modulate gene expression that may be advantageous in adapting to new metabolic landscapes.

3.4 Discussion

Microbial adaptation during colonization is a dynamic process

Bacteroides species are dominant and prevalent in many human populations. Instrumental to their success is their ability not only to tolerate stress, but to quickly adapt at growing in a range of conditions. Bt has been used as a model organism to understand the genetic drivers of gut colonization for commensal organisms (228, 234, 235). However, by and large, a single representative timepoint has been assessed in previous investigations, which fail to take into account the dynamism of the adaptational process. For example, contrary to the established characterization of Bt as exhibiting constant, rapid growth in the gut based on gene expression at a single timepoint (236), our time course data evaluating CFU counts (Fig. 17), transcriptomics, and functional genetics suggest that upon entry into the gut ecosystem Bt growth lags for at least 4 days, during which Bt prioritizes biosynthesis of amino acids and other essential compounds, before it begins a period of rapid growth. In this study, we find that the set of genes required for Bt to initially establish colonization in the gut are distinct from those required to persist (Fig. 8). Though the findings of previous studies do reflect a replicable snapshot of the Bt adaptational process, these studies have largely neglected the ecological dynamics of Bt adaptation that play out temporally on the scale of weeks. The results presented here suggest that investigators should think carefully in selecting the timepoints that are most relevant to their investigation of gut colonization and adaptation.

Dynamically changing pathways in Bt

The process of colonizing a new host presents a collection of environmental changes that require physiological adaptation. For example, the bacteria experience changes in pH and nutrient availability as well as assaults by the host immune system. The largest transcriptomic changes that we observed across the two weeks of Bt colonization occurred in the first week: 10% of all Bt genes (503 genes) were differentially expressed between D1 and D7, whereas only 0.1% of the genes were differentially expressed between D7 and D14 (Fig. 10A).

Our findings suggest Bt faces amino acid starvation upon arrival in the lower GI tract during the earliest stage of colonization. During amino acid starvation, microbes activate a program of transcriptional changes called the stringent response, in which carbohydrate metabolism is curtailed in favor of amino acid biosynthesis. Schofield et al. (237) previously showed that Bt requires the stringent response to establish and maintain colonization in ex-germ-free mice. Indeed, in this study we demonstrated that biosynthesis of amino acids is both highly upregulated and functionally vital to the initial establishment of Bt colonization. One possible explanation for why amino acid biosynthesis pathways have significantly higher expression on D1 compared to D7 or D14 is that when Bt first arrives in the uncolonized gut, dietary or host-derived amino acid resources are scarce enough to induce the stringent response, which leads to the general upregulation of amino acids. Upon adjustment to the environment, it will be more energetically favorable to fine-tune expression of only the necessary resources.

In addition to amino acid biosynthesis, the other major category of genes that were highly expressed on D1 but not D7 or D14 are those involved in vitamin biosynthesis. Genes including those related to biotin (Vitamin B7), riboflavin (Vitamin B2), and pantothenic acid (Vitamin B5)

biosynthesis were significantly enriched on D1. Like amino acids, these vitamins may simply be scarce in the lower GI tract early during colonization of ex-germ-free mice, requiring Bt to synthesize them itself early in colonization. Indeed, dietary B vitamins are primarily absorbed by the host in the small intestine, and although microbe-derived B vitamins can be produced and absorbed in the colon, in our germ-free model, we would expect B vitamin levels to be low in the distal GI tract prior to colonization (238). It is also possible that the upregulation of biotin and other vitamins could be linked to host immunity. Given that microbial biotin biosynthesis genes are upregulated during the infection of many pathogens in humans, including pathogenic E. coli gastrointestinal infections and Salmonella bloodstream infections (239–241) and host biotin deficiency has been demonstrated to increase secretion of proinflammatory cytokines in human dendritic cells (242), upregulation of biotin biosynthesis by Bt may serve to dampen inflammation from the host in response to bacteria colonization. Though Bt is not traditionally regarded as a pathogen, we cannot rule out the possibility that it can induce some level of inflammation, especially in mice that have never encountered microbes. Bacteroides fragilis is a well-studied species of the same genus that is commonly part of a healthy microbiome but can induce inflammation within the host under specific contexts (18, 19). Future investigation will be required to understand whether upregulation of vitamins in Bt is related to any of these host responses.

Loss of diversity in RB-Tn pools in vivo and in vitro and the impact of complex communities

In both the mouse gut and in complex broth (BHI-S), our RB-Tn mutant pool exhibited significant losses in Shannon diversity during the first week, which extended through the remainder of the two-week experimental period (Fig. 18C; Fig. 19). Whereas in vitro loss of diversity was primarily attributable to shifts in the relative abundance of the mutants rather than extinction of mutant strains, the collapse in diversity observed in the host-associated RB-Tn pool was caused by decreases in both strain evenness and richness—and on a much larger scale than in the in vitro experiments. This difference in scale suggests that *in vivo, strong*, directed selective pressure for improved melibiose and RFO metabolism led to population sweeps by strains carrying gain-of-function mutations. By contrast, the loss of diversity exhibited by our in vitro experiments suggests a much weaker selective environment, which is not surprising given that the mutant population was generated and has been cultivated in vitro in BHI-S. Moreover, WT Bt grows efficiently in BHI-S thus there is limited dynamic range to respond to the selective pressures of that environment with improved growth. By contrast, WT Bt growth on melibiose, is highly inefficient, with a doubling time of around 10 hr, and the RB-Tn gain-of-function mutations were able to increase growth speed on melibiose by log-fold differences.

Like our study, a transposon screen for enteric colonization factors in Klebsiella pneumoniae also reports that despite robust colonization of the colon by a diverse array of mutants at D1, the diversity of the mutant population was greatly reduced by D4 and there was marked expansion of a small subset of mutants (20). However, contrary to our findings, diversity of the mutant Klebsiella population recovered shortly thereafter. This difference can potentially be accounted for by the fact that the Klebsiella experiments were carried out in SPF mice treated with antibiotics that reduced, but did not eradicate, the native microbiome to allow for dense Klebsiella colonization.

We find that, in vitro, addition of an SPF community prevented the loss of diversity in passaging the Bt RB-Tn library for 14 days (Fig. 19). In accordance with these observations, a recent publication investigating the impact of community complexity on evolution of focal strains concluded that microbes have a higher evolutionary capacity in low diversity communities than in complex communities (21). Thus, the strong, directed selection observed in our in vivo experiments may be rendered more diffuse in the presence of a native microbial community.

Among other factors, we expect that the results of this study will likely differ depending on the presence or absence of a pre-established community. Though we recognize the limitations of studying adaptation and colonization in GF mice, these experiments lay foundational groundwork for the eventual investigation of community assembly and fitness dynamics in more complex communities. Specifically, having a thorough catalog of host-specific colonization dynamics in the GF gut will create a basis of comparison from which we can delineate host- versus microbiome-induced selective pressures and adaptive dynamics. Preliminary assessments of RB-Tn mutant engraftment in complex communities revealed that a strong initial bottleneck in colonization led to very low levels of Bt persistence in the community (Fig.17). Moreover, the final composition of mutant strains appeared to be dictated by stochastic forces, with high inter-individual variability, and no consistent selective pattern for particular genes or functions (Fig. 20). This further supports the notion that addition of a complex community may limit the extent of strong, directed selection, although additional follow-up experiments are required to further characterize these patterns.

IS3 transposable elements: a potential novel mechanism to modulate expression of specific CAZymes?

In GF mice fed a standard RFO-rich diet, the selective pressure for mutants with increased alpha-galactosidase activity is severe. Mutants in the BT1871 locus were consistently selected for in WT Bt allowed to evolve in mice for six weeks. In all three populations, the downstream side of the duplicated region ends midway through a ribosomal gene, meaning that there are no known transcriptional terminators between the ribosomal promoter and the alpha-galactosidase gene, BT1871. It is likely that these mutants have increased alpha-galactosidase activity due to readthrough from the strong ribosomal promoter in the upstream copy of the locus. Though increased alpha-galactosidase could be simply due to the presence of two copies of BT1871, the growth phenotype that these mutants exhibit, in which log phase growth is 5x faster than WT, is on par with that of the RB-Tn mutants, for which alpha-galactosidase expression was more than 10 times greater than that of WT.

In Bt identical copies of this IS3 transposable element occurs in seven locations, often adjacent to PULs and almost always paired with a ribosomal gene. It is well known in both mammals and prokaryotes that transposable elements can modulate the expression of nearby and distant genes. Transposable elements may be selected for because the genetic "cargo" that they shuttle along (such as an antibiotic cassette) is beneficial to the organism or because the position where the insertion occurs results in some downstream effect that is beneficial to fitness. Both reasons may contribute to the selection of the tandem duplication of the BT1871 locus. Future studies will be necessary to investigate whether other IS3 elements in Bt perform similar functions in modulating expression of nearby genes and whether such a mechanism can be found in other families of bacteria.

3.5 Materials and Methods

Mice

Female mice 8-12 wk-old C57Bl/6J germ-free (GF) mice were bred and maintained in plastic gnotobiotic isolators or bioexclusion racks within the University of Chicago Gnotobiotic Core Facility and fed ad libitum autoclaved standard chow diet (LabDiets 5K67). The mice were given a single dose of either wildtype Bt VPI-5482 or the Bt mutant library (gift from Deutschbauer lab) at 106-108 CFU / 200µL. All murine experimental procedures were institutionally approved.

In vivo genome-wide mutant fitness assays

Several individual cohorts of mice were used, indicated by the month of the experiment. For each cohort, mice were housed in a single gnotobiotic isolator for the duration of the experiment (December and January), or each individual cage was housed in a bioexclusion rack (March and October). Female C57BL/6J GF mice between 8-12 weeks old were fed a standard irradiated diet ad libitum, split into cages of n = 2-3 mice/cage, and allowed to acclimate for 3 days prior to colonization. The inoculum was prepared by one of two methods: 1) for Mar. and Oct. experiments: thawing a 2mL aliquot of the Bt RB-TnSeq library, growing the entire aliquot in 150mL BHI-S medium overnight with 20 ug/mL erythromycin (16h), and backdiluting the culture to OD600=0.05 the next morning to allow for cells to reach mid-log phase (3hr), or 2) for Dec. and Jan. experiments: thawing aliquots of the Bt RB-TnSeq library and gavaging the thawed Bt cells directly into mice. For each experiment, at least 3 cell pellets of the inoculum were collected at Time=0 references. Each mouse was colonized by oral gavage with 200µL of

the Bt transposon library. Stool samples were collected daily (excluding weekends), up to 14 days post-colonization to assess longitudinal shifts in mutant abundance. Mice were monitored and weighed daily. Genomic DNA was extracted using the DNeasy PowerSoil Kit.

Pipeline for measuring relative abundance and fitness scores of RB-Tn mutants

RB-TnSeq strain and gene fitness scores were calculated as described previously from strain-level count data (322). For temporal abundance analyses of TnSeq mutants for each run, we first created a feature table with raw counts of strain-level mutants across each sample. This table was filtered to remove strains with counts of 1, as these are likely produced by sequencing error. Next, counts were normalized by the count of a synthetic spike-in barcode that was introduced at 20pM into each sample during PCR amplification of the barcodes. Samples where the spike-in represented > 30% of total reads were discarded. The synthetic spike-in barcode was subsequently removed as a feature from the table, and the resulting tables were used for alpha diversity analyses via the R package vegan. For all other relative abundance analyses, strains were assigned to genes based on previous mapping by Liu et al. 2021 (19), as well as manual mapping performed for this experiment. Strains that had not been previously mapped were binned together as "non-mapping" strains, and strain-level counts were then summed for each gene. For all subsequent analyses, we further filtered out genes that mapped to Bt plasmids, as we were more interested in chromosomal gene fitness patterns. The R package phyloseq was used to calculate Bray-Curtis dissimilarity for all pairwise combinations of samples, which was then used to create PCoA plots. Finally, filtered count tables were adjusted to relative abundance based on the total remaining counts, and used to track gene-level relative abundance over time.

For linear regression of initial vs final relative abundance of gene mutants, genes with zero-counts were re-assigned a value of $1e^{-7}$ to perform log-transformation of the data.

In vivo transcriptomic experiments

Mice were sacrificed at 1, 7, and 14 days post-colonization. Luminal contents of the mice cecum were immediately snap-frozen in liquid N2 and stored at -80C. About 50mg of the contents were transferred into 2mL screw-cap tubes, followed by the addition of 1mL TriZOL reagent for the isolation of RNA. The samples were homogenized by beadbeating with 0.1-mm glass beads in a Mini-BeadBeater-96 for 2 minutes. Total RNA isolation and purification were performed using the TRI reagent protocol and quality checked by BioAnalyzer. All library preparation and sequencing work was performed by MiGS (Pittsburgh, PA). Initial DNAse treatment is performed with Invitrogen DNAse (RNAse free). Library preparation is performed using Illumina's Stranded Total RNA Prep Ligation with Ribo-Zero plus. Custom Ribo-Zero probes were designed for Bt and supplemented alongside the standard probe set. Custom probe sequences can be found in Table 8. Sequencing was performed on a NextSeq2000 giving 2x50bp reads. Post sequencing, we use bcl2fastq (v2.20.0.422) to demultiplex and trim adaptors.

Gene Set Enrichment Analysis (GSEA) of metabolic pathways

Metabolism for the Bt genome was first estimated using the anvio-v7 program "*anvi-estimate-genome"* that identifies the KEGG Ortholog family (KOfam) annotations for each open reading frame. Gene calls for each metabolic pathway found within the B. theta genomes were then transferred into a unique GSEA pathway query list in R (*'fgsea')*. Pathway enrichment was then

calculated using the enriched gene lists derived from the RNA-Seq analyses using '*deseq2*'. Pathways were filtered for *padj>0.05* and the normalized enrichment scores (NES) plotted (PRISMv9). Other custom gene lists were created to calculate the enrichment of other gene sets including the polysaccharide utilization loci (PULs), genes involved the production of specific amino acids, and capsular polysaccharide loci (CPSs).

Isolation of mutants from fecal matter

Mouse feces were collected and immediately homogenized in 500mL 25% glycerol solution and stored at -80C. Prior to isolation, glycerol stocks were allowed to thaw on the benchtop for 10 minutes, spun for 30 seconds at 2,000 RPM. On each 150mm BHI-S plate, 100uL of 10-3, 10-4, or 10-5 dilution of the glycerol stock was spread using 4.5 mm glass beads. The plates were incubated anaerobically at 37C for two days. Individual colonies were picked into 1mL 96-well plates containing 750uL BHI-S in each well. After 16 hours of growth, glycerol was added to a final concentration of 20% and the isolates were stored. The isolate stocks were used as the template in PCR amplifying the barcoded region of the mutants. The PCR products were sent for Sanger sequencing.

***In vitro* culture and growth measurements**

Bt was grown in an anaerobic chamber at 37°C either in Brain Heart Infusion Supplemented (BHI-S) medium or Varel-Bryant (VB) defined medium. The Varel-Bryant medium base was made with no carbon source and was supplemented by 20mM of the indicated carbon source. For

growth measurements, colonies of Bt were inoculated into 3mL of BHI-S in plastic culture tubes and grown overnight at 37°C, for a total of 6 biological replicates. Sealed Hungate tubes containing 10mL of VB medium and a 20mL headspace were used for subsequent growth. Immediately before inoculating with the cells, the tubes were inoculated via syringe with autoclaved sugar solution and hemin solution for a final concentration of 20mM sugar and 5 ug/mL hemin. Overnight cultures were diluted to OD600 = 1, and 100uL of the diluted culture was inoculated into the prepared media tubes via syringe. Anaerobically sealed cultures were grown outside of the chamber in a 37°C incubator with no shaking. Every 45 minutes, the cultures were taken from the incubator, cells resuspended by shaking, and their OD600 readings measured by a GENSYS 40-Vis spectrophotometer.

Host-associated evolution of spontaneous Bt mutants

Three female mice 8-12 wk-old C57Bl/6J GF mice were co-housed in the same gnotobiotic isolator and fed standard chow diet ad libitum. One week post inoculation of Bt, the three mice were separated into individual cages. A fecal sample was taken six weeks post inoculation. DNA was extracted from the fecal pellets by the phenol-chloroform method, followed by ethanol precipitation, and sent for shotgun sequencing. Individual isolates from each fecal pellet were cultured from the bulk material as outlined in Isolation of mutants from fecal matter. We assayed for isolates with increased growth in VB-melibiose and selected one isolate from each mouse for MinION long-read sequencing.

Isolated genomes sequencing, assembly and polishing

To provide greater context for the delineation of the complex chromosomal rearrangements associated with the BT1872/BT1873 operon, a long-read sequencing strategy was employed. The isolate genomes assessed were wild-type Bt, the strain used for the mouse experiments, and three spontaneous mutant cultivars (MZ55, MZ58 and MZ65), which demonstrated enhanced growth rates in the presence of melibiose, recovered from the feces of mice six weeks after initial inoculation. Total genomic HMW DNA was extracted by a standard phenol chloroform protocol (246) on overnight 25-mL BHIS broth cultures. DNA was resuspended in 0.1-mL 10 mM Tris-Cl, pH 8.5.

Slow pipetting, wide bore pipette tips and steps to minimize velocity gradients were implemented throughout to avoid further shearing of DNA molecules. Libraries were prepared with the Rapid Barcoding Kit (SQK-RBK004) and the standard protocols from Oxford Nanopore Technologies were used with the following modifications. DNA fragmentation was performed on 10-ug DNA using 10 passes through a 22G needle in a 250-μL volume before purification using 0.5% Agencourt AMPure XP beads (A63882, Beckman Coulter). Each elution step of the AMPureXP beads was performed using 10 mM Tris-Cl pH 8.5 instead of water, at 37°C for 5 min. The gDNA inputs into library preparation ranged between 0.5 μg and 1.2 μg (Table S2), based on sample availability in a standard 8.5 ul volume, with 1.5 μl Fragmentation mix added to each sample. Barcoded libraries were pooled so each sample contributed an equal input mass (~0.5 μg, Table 9). Using MinKNOW (v4.3.4), a single R9.4/FLO-MIN106 flow cell (Oxford Nanopore Technologies) sequenced the final prepared library with a starting voltage of –180 mV and a run time of 72h. Guppy (v5.0.11) and the sup model were used for post-run basecalling, sample de-multiplexing and the conversion of raw FAST5 files to FASTQ files. For downstream

analyses, we only used reads with a minimum quality score of 7. We assembled long-reads contigs with Flye (247). Additional DNA extractions were carried out for every isolate using a standard phenol-chloroform extraction and send for short-read sequencing. We then used the short-reads to polish the long-read assemblies using Pilon v1.23 (37).

Metagenomic mapping and coverage visualization

We used anvi'o v7.1 (249) and the metagenomic workflow to compute and visualize metagenomics coverage for each isolate genome. Briefly, the workflow uses (1) Prodigal v2.6.3 (250) to identify open-reading frames (ORFs), (2) 'anvi-run-hmm' to identify single copy core genes from bacteria (n=71, (40) and ribosomal RNAs (n=12, modified from https://github.com/tseemann/barrnap) using HMMER v3.3. (252), (3) 'anvi-run-ncbi-cogs' and 'anvi-run-kegg-kofams' to annotate ORFs with the NCBI's Clusters of Orthologous Groups (COGs) (253) and the KOfam HMM database of KEGG orthologs (KOs) (254, 255) respectively. We used Bowtie2 v2.3.5.1 (45) to recruit metagenomic short-reads to the contigs, and samtools v1.11 (257) to convert SAM files to BAM files. We profiled the resulting BAM files with 'anvi-profile' and used the program 'anvi-merge' to combine all single profiles into a merged profile for downstream visualization. We used 'anvi-get-split-coverages' and 'anvi-script-visualize-split-coverages' to generate the coverage plots. We used 'anvi-export-gene-calls' and gggenes v0.4.1 to visualize the genomics context around BT1871.

BT1871 copy number

We used blast (258) to compute the number of long-reads with two copies of the BT1871 locus in the MZ65 isolate. We extracted the gene sequence (1989 bp) from the initial B. theta gnome (AMD595) using anvi'o interactive interface and used blastn v2.5.0 to blast the long-reads from MZ65. Blast hits with an alignment length > 180% of the gene length were flagged as "two copies" and hits with alignment length between 80% and 105% were flagged as "one copy". To visualize long-reads with two copies of BT1871, we used minimap2 v2.17 (259) to map the MZ65 long-reads to the MZ65 genome, which had two copies of the BT1871 region. We used samtools v1.11 to extract the reads with two copies as identified above and used IGV v2.11.1 (260) to visualize the mapping and generate a figure.

Metabolite Extraction from Cecal Material

Metabolites were extracted with the addition of extraction solvent (80% methanol spiked with internal standards and stored at -80°C, Table 10) to pre-weighed fecal/cecal samples at a ratio of 100 mg of material per mL of extraction solvent in beadruptor tubes (Fisherbrand; 15-340-154). Samples were homogenized at 4°C on a Bead Mill 24 Homogenizer (Fisher; 15-340-163), set at 1.6 m/s with 6 thirty-second cycles, 5 seconds off per cycle. Samples were then centrifuged at -10°C, 20,000 x g for 15 min and the supernatant was used for subsequent metabolomic analysis.

Metabolite Analysis using GC-EI-MS and Methoxyamine and TMS Derivatization

Metabolites were analyzed using GCMS with Electron Impact Ionization. 100uL of metabolite extract was added to prelabeled mass spec autosampler vials (Microliter; 09-1200) and dried down completely under nitrogen stream at 30 L/min (top) 1 L/min (bottom) at 30°C (Biotage SPE Dry 96 Dual; 3579M). To dried samples, 50 μL of freshly prepared 20 mg/mL methoxyamine (Sigma; 226904) in pyridine (Sigma; 270970) was added and incubated in a thermomixer C (Eppendorf) for 90 min at 30°C and 1400 rpm. After samples are cooled to room temperature, 80 μL of derivatizing reagent (BSTFA + 1% TMCS; Sigma; B-023) and 70 μL of Ethyl Acetate (Sigma; 439169) were added and samples were incubated in a thermomixer at 70°C for 1 hour and 1400rpm. Samples were cooled to RT and 400 μL of Ethyl Acetate was added to dilute samples. Turbid samples were transferred to microcentrifuge tubes and centrifuged at 4°C, 20,000 x g for 15 min. Supernatants were then added to mass spec vials for GCMS analysis. Samples were analyzed using a GC-MS (Agilent 7890A GC system, Agilent 5975C MS detector) operating in electron impact ionization mode, using a HP-5MSUI column (30 m x 0.25 mm, 0.25 μm; Agilent Technologies 19091S-433UI) and 1 μL injection. Oven ramp parameters: 1 min hold at 60°C, 16°C per min up to 300°C with a 7 min hold at 300°C. Inlet temperature was 280°C and transfer line was 300°C. Data analysis was performed using MassHunter Quantitative Analysis software (version B.10, Agilent Technologies) and confirmed by comparison to authentic standards. Normalized peak areas were calculated by dividing raw peak areas of targeted analytes by averaged raw peak areas of internal standards.

Chapter 4: Conclusions, Challenges, and Future Directions

4.1 Preface

Our work studying genetic adaptations of two different model organisms in the context of host-associated selective pressures has advanced our understanding how these bacteria adapt to environmental transitions. In Chapters 2 and 3, I detailed our discoveries with regards to identifying antagonistic gene-environment interactions within an antibiotic context and assessing the genetic determinants of a major commensal microbe colonizing a mammalian gut. Here, I will discuss the major conclusions, challenges, and prospective for future studies.

4.2 Contributions to method development

In addition to the findings of the above studies, we also made significant progress toward optimizing the pipeline for these discoveries.

4.2.1 Systematic approach to identify extreme antagonistic interactions

For the E. coli study, we were interested in studying extreme antagonistic interactions. In the current literature, extreme antagonistic literature pairs seem to be quite rare, and there lacked a systematic method for identifying such interactions other than testing existing antibiotics in pairwise fashion (207). The methodology of our study involves serial adaptive evolution of E. coli by first adapting the bug to a defined glucose medium, isolating mutants, then adapting those strains to a sublethal concentration of rifampicin. We chose rifampicin because there is not yet an antagonistic partner antibiotic for this drug that is still used for the treatment of serious bacterial pathogens such as Mycoplasma tuberculosis.

Currently, antibiotic cycling strategies are implemented on an ad hoc basis, and the effectiveness of such a regime is still inconsistent and controversial (188, 327–329). However, if we were to implement this regime based on knowledge that resistance to one environmental perturbation (Antibiotic 1) increased sensitivity to second environmental perturbation (Antibiotic 2), then we could leverage this collateral sensitivity and limit expansion of the resistant strains by cycling between the extreme antagonistic pair. In theory, one of the antibiotics will be targeted toward suppressing the growth of the pathogen, and the other will be used to suppress resistant mutants of the first. An important arm of our strategy is preadapting the organism to the "base"

environment such that the adaptive mutations will not also enhance fitness in the absence of an antibiotic. In other words, these mutants are sensitized to the addition of rifampicin. Therefore, the mutations identified in the second adaptive step constitute candidate for the development of antagonistic partners for rifampicin. This is an improvement to the current strategy of selecting pairs from existing drugs because we can design and optimize drug pairs for this regime without being limited to drugs currently on the market. Finally, this approach is scalable to identifying higher order interactions (i.e., antagonism between three or more environmental perturbations), and can be further optimized by using mutagenesis methods to accelerate generation of variability during the process of target identification.

4.2.2 Three-prong approach for assessing genetic determinants of colonization

Colonization is a complex process involving inputs from the host, microbes, and abiotic factors such as the food being ingested. Severe cut-offs are applied to transcriptomic datasets in differential gene expression analysis because there are many sources of biological and technical noise and problems inherent in statistical models of count data (330). This is also true for calculation of BarSeq fitness scores (331).

To more comprehensively assess whether the genetic determinants of gut colonization shift over time, we combined two complementary approaches: transcriptomics (RNA-Seq) and functional genetics (RB-TnSeq). These are both unbiased methods for identifying genes most important to the fitness of a given microbe under a particular set of conditions. Bacterial transcriptomics offers valuable information about gene usage, allowing us to identify differences in gene expression at progressive timepoints after introduction, and to evaluate the functional annotations

of those genes. A limitation of this approach is that conditional (or temporal) differences in gene importance may not be reflected in changes at the transcriptional level. Furthermore, there are examples in which modest transcriptional changes may have major functional consequences (332), as well as many examples in which post-transcriptional regulation determine levels of functional protein (333–335).

These are caveats that can be addressed by a functional genetics assay, RB-TnSeq followed by BarSeq (300, 331), in which we allow a diverse pool of mutants to compete with one another and track their relative abundance at different points in time. Mutants that have disruptions in important genes will be depleted from the population, whereas mutants that retain the important genes for fitness or carry mutations that further confer further fitness advantages will be enriched over time. Therefore, this assay allows us to make conclusions about the functional consequences of each genetic disruption. A clear limitation of this approach is that we will not be able to identify genes with vital but redundant functions but given that there are significant differences in their expression, these genes should still be identified by the transcriptomics approach.

Invariably, sole reliance on a single method, whether it be differential expression or functional genetics, will result in false positives as well as false negatives in identifying genes critical for survival in each condition (328). All candidate genes selected from these datasets should be verified through a series of secondary assays *in vitro* and *in vivo*, but these secondary assays are often time-consuming and lacking in throughput.

Another method that we used to further assess the validity of our genetic hits is to measure the metabolite levels within the cecal contents of our mice. By measuring a panel of amino acids and

carbohydrates, we were able to validate our findings regarding the upregulation and functional importance of amino acid biosynthesis during initial introduction of Bt and its subsequent shift toward prioritizing carbohydrate metabolism midway through the first week and onwards. Furthermore, data from transcriptomics and functional genetics guided our assessment of the appropriate metabolites, which is valuable considering that a major challenge in designing metabolomics experiments is that important differences may not be detected if those specific compounds are not on the common panel of metabolites run by the facility.

Together these approaches provide complementary insights into genetic requirements through time. This method for identifying genetic adaptations to colonization can be extended to various models of disease and serve as a basis upon which exploration of more complex microbial communities will build.

4.3 The stringent response for microbes in environmental transition

We used publicly available data comparing serine hydroxamate treatment between WT *E. coli* K12 and its *relA*-deficient mutant to identify the transcriptional signature of the stringent response. The *E. coli* stringent response downregulates genes involved in ribosome synthesis and upregulated those involved in amino acid biosynthesis in response to starvation. We find that the fastest-growing mutants have transcriptional program resembling the stringent response in the absence of rifampicin, whereas slower strains do not exhibit this response. This suggests that the fast-growing mutants have evolved to grow more efficiently and more narrowly in a certain environmental condition (i.e., in the presence of rif) such that when the condition is altered, the environment becomes metabolically challenging to the organism and elicits activation of the stringent response. Indeed, the metabolic program of these mutants are anticorrelated with those of the slow-growing mutants (Fig. 6). This result is also interesting in the context of the findings of our colonization study (Chapter 3). Transcriptomics, functional genes, and metabolomics all suggest the presence of a stringent response-like regime in which carbohydrate metabolism is temporarily suppressed in favor of higher amino acid biosynthesis.

4.4 Colonization of commensal microbes in the presence of a defined community

In our study using GF mice as a model to investigate the genetic requirements of a commensal organism to colonize a murine gut, we characterize gut colonization as dynamic process consisting of multiple stages with shifting genetic determinants of microbial fitness. We not only identify key pathways involved in colonization but determine the timing of *when* these pathways will be most vital to colonization success.

To improve microbiome-based therapies, it will be important to appreciate and meet the distinct needs of the organism during each stage of colonization. Finally, the results of this study contribute to a growing need to understand adaptative processes in microbes, and the complex biotic and abiotic forces which guide them. Future studies will expand on these findings in different ecological contexts and host backgrounds and explore the effect of various microbial communities on these dynamic genetic requirements.

4.5 Evaluating colonization determinants with a pre-established community

We recognize the limitations of studying adaptation and colonization within a GF mouse model. Through preliminary testing of our pipeline within a SPF mouse model, we recognize several potential challenges. Firstly, when we introduced the RB-Tn mutants into a pre-established complex community, we observed a strong initial bottleneck in colonization that led to low levels of Bt persistence in the community (Fig. 17). Second, perhaps partly because of this bottleneck, the composition of the mutant strains within the SPF model appeared to be dictated at least in part by stochastic forces with high variability within each mouse (Fig. 20). Unlike the severe uni-directional selection that occurred within the RB-Tn population in GF mice, there was no consistent selective pattern for specific genes or function for the same library when there was a pre-established community. This may suggest that the presence of a complex community limited the extent of adaptation that could occur within a single population of the community. When a community is sufficiently diverse, even if the selective pressure is strong overall within an environment, the population in question may not have sufficient access to the resources that maximize incentive to evolve, the real selective pressure felt by that population might be lower (184). Additionally, in a complex community that efficiently utilizes resources, there will not be a dominant nutrient available in excess. Furthermore, in the context of entering a pre-established community, available niches are much more limited compared to colonizing of a GF gut; therefore, mutants that are best adapted to persist in the gut may be lost to the initial bottleneck, rendering the mutants in the rare penetrating population effectively neutral in term of fitness.

To circumvent the problem of low engraftment in a pre-established community, antibiotics have been used as a pretreatment before the organism of interest is introduced (302, 329). This is

appropriate in contexts when investigating organisms that are given as an FMT, which are often administered after antibiotic treatment, but does not help to answer questions of how exogenous organisms enter a pre-established community that has not recently received a serious perturbation. Therefore, pre-treatment of antibiotics is not always possible as a workaround for engrafting an organism of interest into a complex community. In fact, while thinking about how to improve engraftment of probiotics, we should consider the biological forces that exclude or strongly restrict new exogenous organisms from engraftment. Rather than a large perturbation such antibiotics, is it possible to temporarily create a new niche for the exogenous strain such that the population can penetrate the gut in large enough numbers to allow for quick and efficient adaptation within the new environment? One such solution was proposed in a recent study, which generated an exclusive metabolic niche for engraftment of an exogenous *Bacteroides* strain that harbored a rare gene cluster that allowed for metabolism of the marine polysaccharide porphyran (337). Strategies such as these open the possibility for engraftment of exogenous strains into a pre-established community without annihilation of the density and diversity of the gut microbiome, which can make the host susceptible to many health risks.

Appendix A – Figures

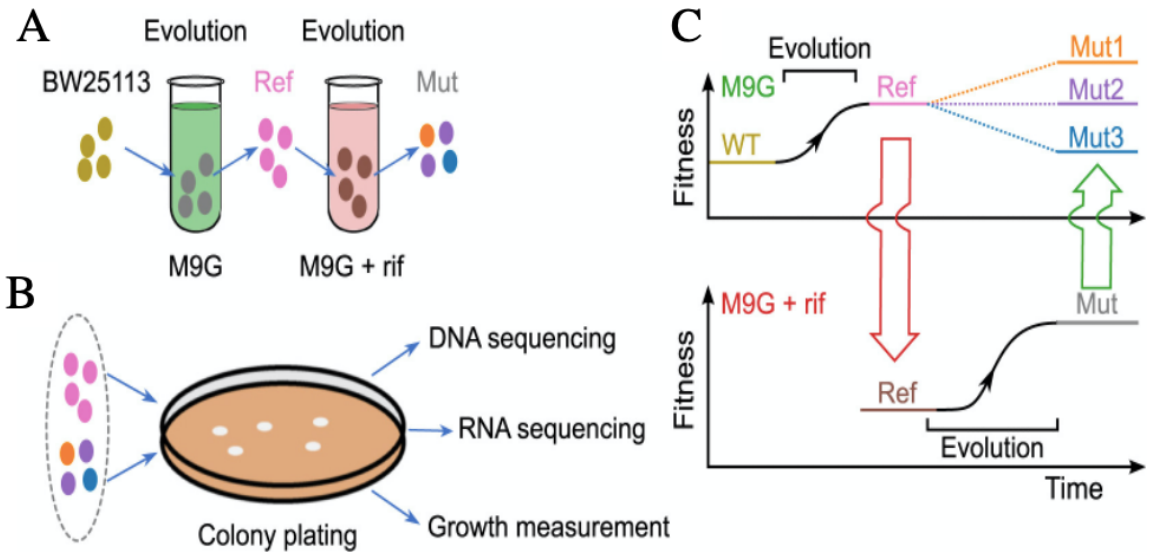

Figure 1: Experimental approach to identify extreme antagonistic interactions.

(A) Diagram of our experimental evolution approach, in which WT E. coli BW25113 was adaptively evolved in M9G, yielding two Ref strains with an enhanced growth rate. These Ref strains were then evolved in M9G + rif, yielding five rif-adapted Mut strains per Ref parent.

(B) Analysis of rif-adapted Mut strains and their Ref parents. Selected colonies were analyzed for growth in the presence and absence of rif, were subjected to WGS to identify mutations associated with adaptation and were RNA-sequenced to detect changes in transcript abundance.

(C) Fitness of the mutants at each stage of the workflow, as measured by growth rate. The growth rate of the rif-adapted strains was compared with that of their Ref parents: Mut strains with significantly slower growth rate in the absence of rif meet the conditions for extreme antagonistic interactions with the addition of rif (Mut3, bottom), whereas Mut strains with growth rate equal (Mut2, middle) or greater (Mut1, top) do not.

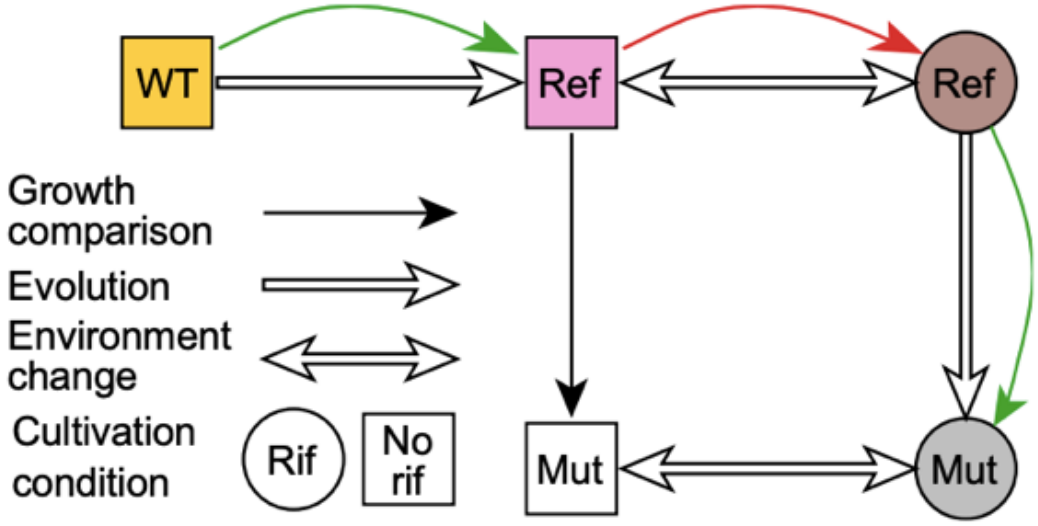

Figure 2: Schematic of the assay to detect extreme antagonistic interactions.

The WT strains were adaptively evolved in M9G medium, resulting in Ref strains. The Ref strains were shifted to an antibiotic medium condition, which was treatment with 7.5 μg/mL rif. Fast-growing, rif adapted, mutants were subsequently evolved, yielding Mut strains. The growth rates of the Ref and Mut strains were measured in both the original (M9G) and stressful (M9G + rif) media. In the diagram, evolutionary changes are represented by single-headed arrows and environmental shifts are represented by double-headed arrows. Green and red edges indicate increases and decreases in growth rate, respectively. The node colors match their corresponding strains in Fig. 1C.

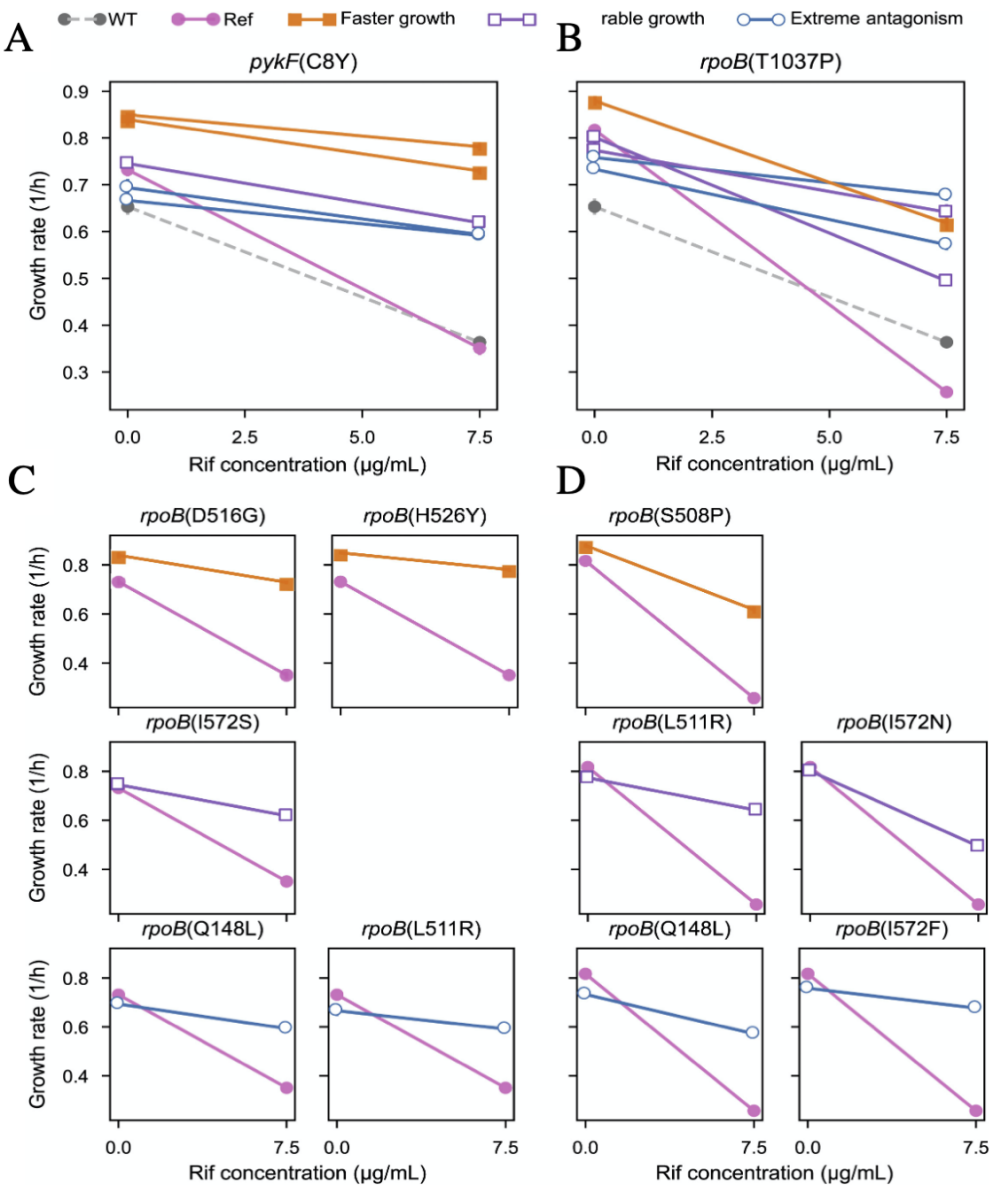

Figure 3: Growth rates of the WT, Ref, and Mut strains in the presence and absence of rif.

(A and B) Growth rates of the WT strain, Ref strains, and Mut strains derived from the pykF(C8Y) parent (A) and the rpoB(T1037P) parent (B). Solid squares, open squares, and open circles indicate Mut strains that grow faster, as fast as, and slower than the Ref strain in the absence of rif, respectively.

(C and D) Growth rate comparisons of the Mut strains with the pykF(C8Y) parent (C) and the rpoB(T1037P) parent (D) in the presence and absence of rif. The means are denoted by symbols and the standard errors of the means are denoted by error bars, which are smaller than the size of the symbol in most cases.

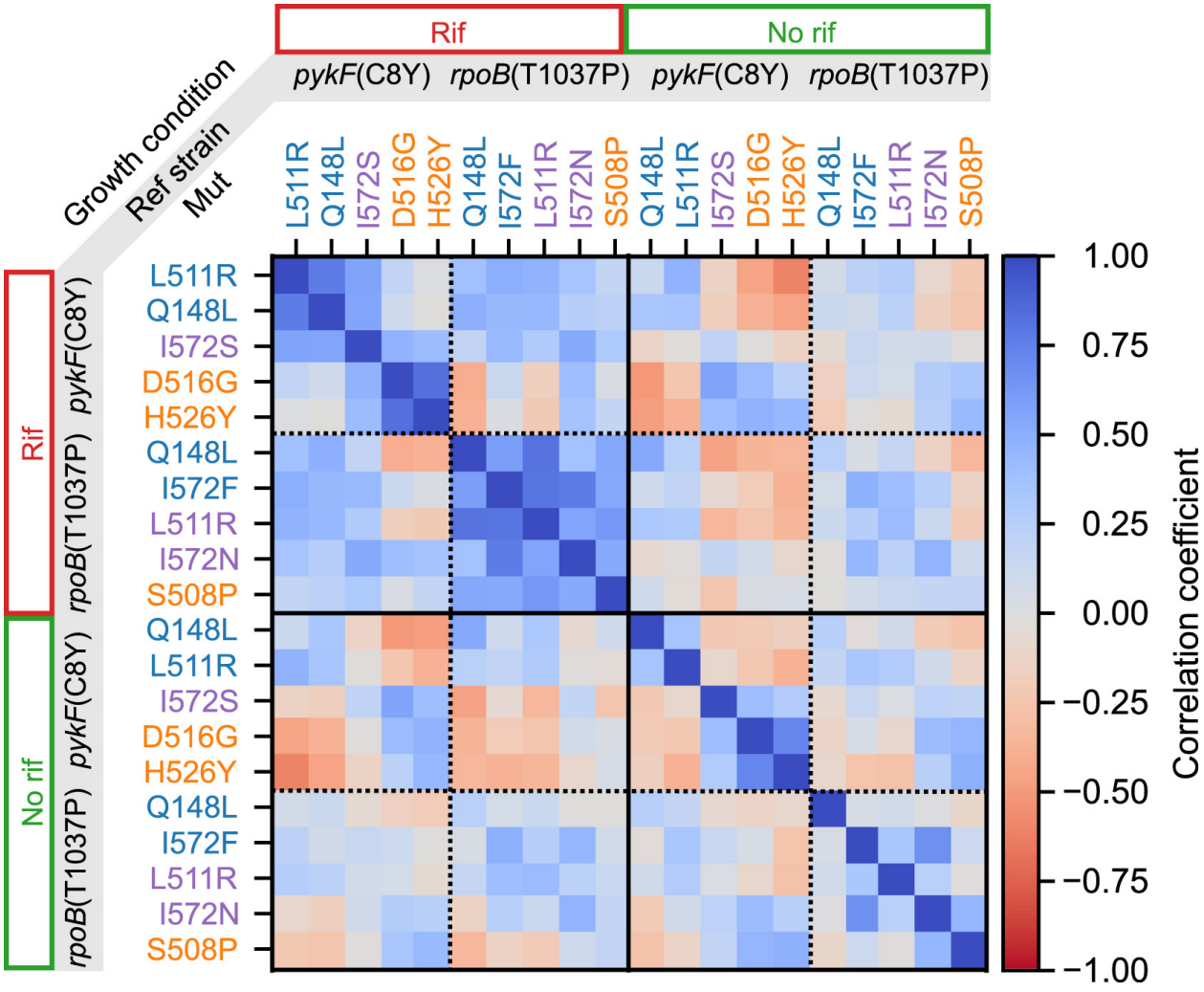

Figure 4: Transcriptional correlations between the Mut strains in M9G and M9G + rif.

The color code indicates the correlations between log2 fold changes for each strain in each cultivation condition. The colors of the labels indicate the growth of each Mut strain relative to its Ref parent as defined in Fig. 1.

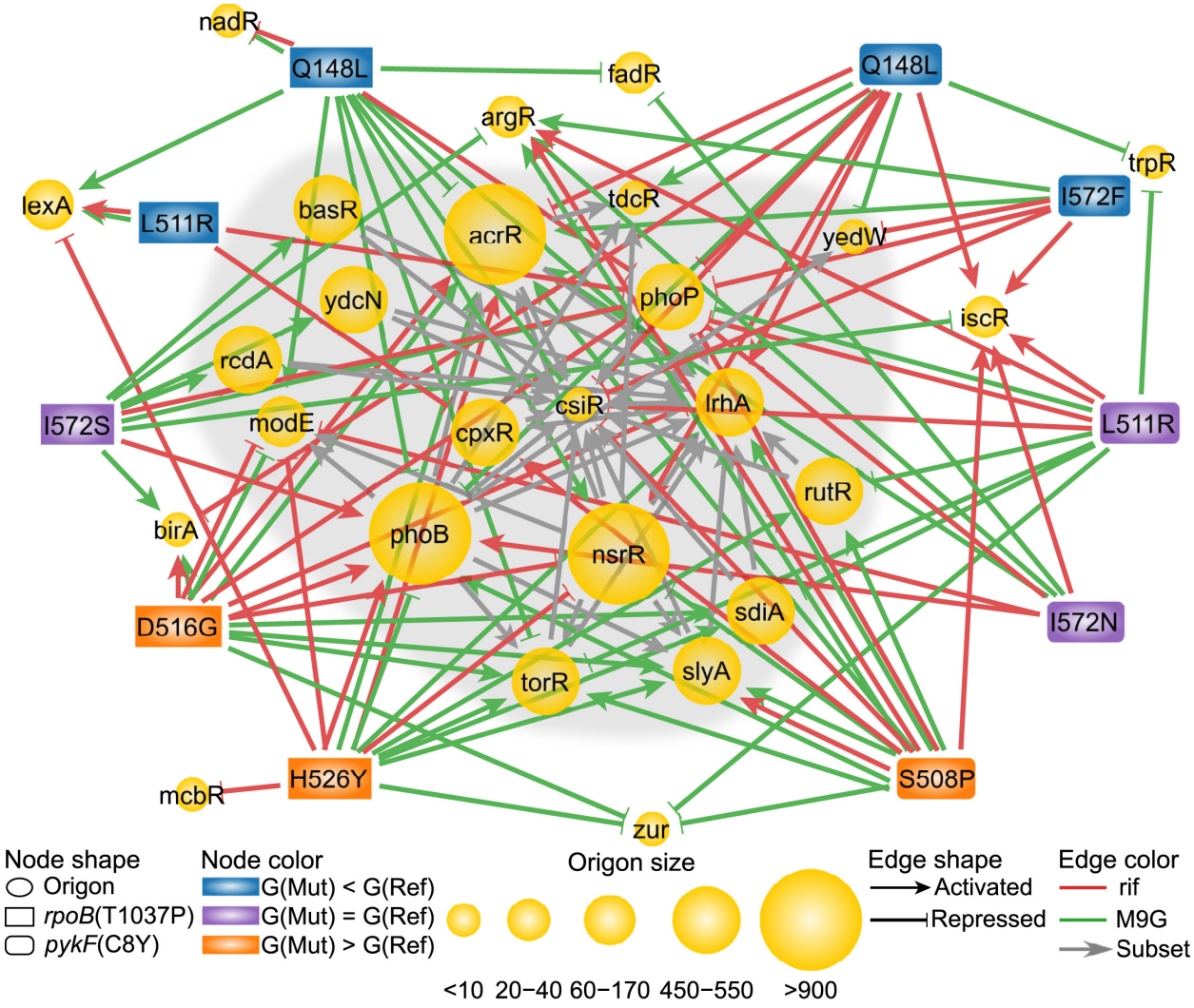

Figure 5: Network of significant origon expression changes in the Mut strains relative to Ref parents.

Network of activated and repressed transcriptional factors across each Mut genetic background in both M9G and rif. The node color, node shape, node size, edge color, and arrow shape are defined in the legend, in which G(Mut) denotes the growth rate of the Mut strain in M9G. Origons are labeled by their master regulator, and all origons that have a subset relationship with another are placed on the gray background.

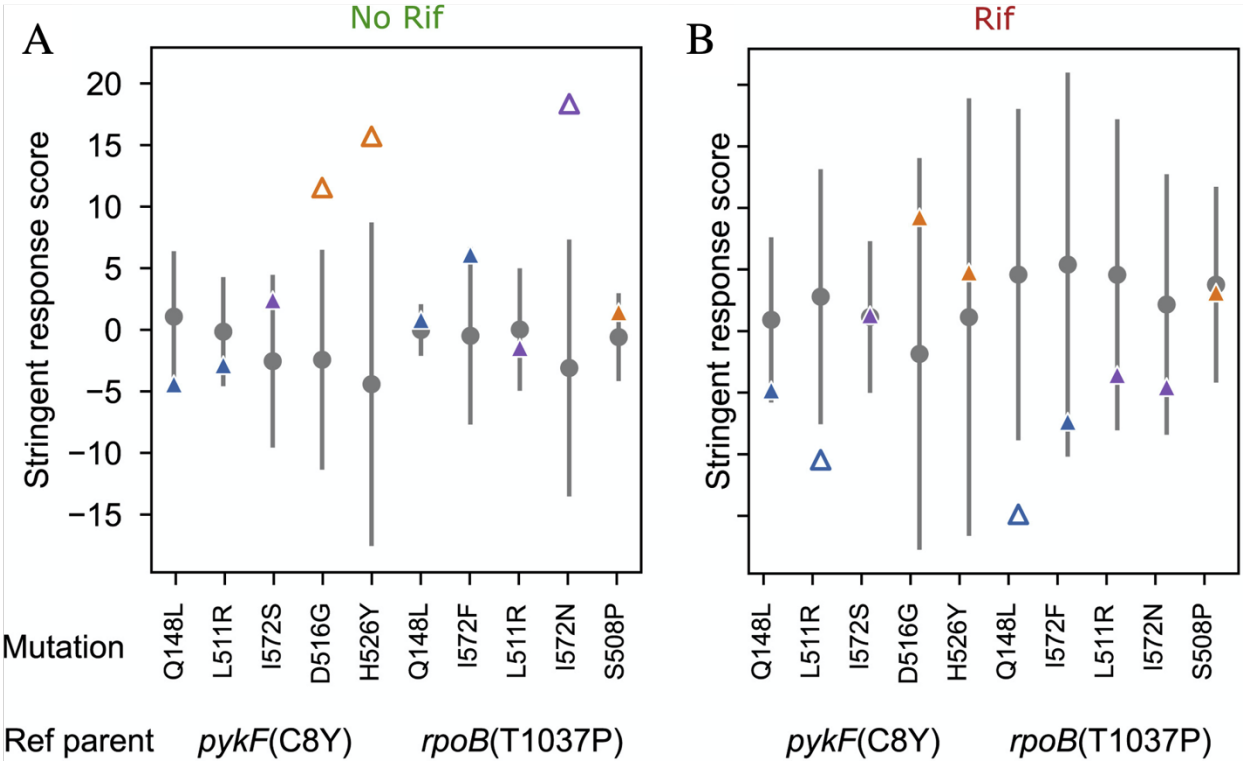

Figure 6: Stringent response score for each Mut strain.

Comparison between the observed stringent response scores (triangles) and bootstrapped means (circles) in M9G (A) and M9G + rif (B). Statistically significant cases are denoted by the open symbols, where the error bars correspond to the 95% confidence interval determined by bootstrapping. The color code indicates the Mut strain's growth rate relative to the Ref parent as defined in Fig. 1.

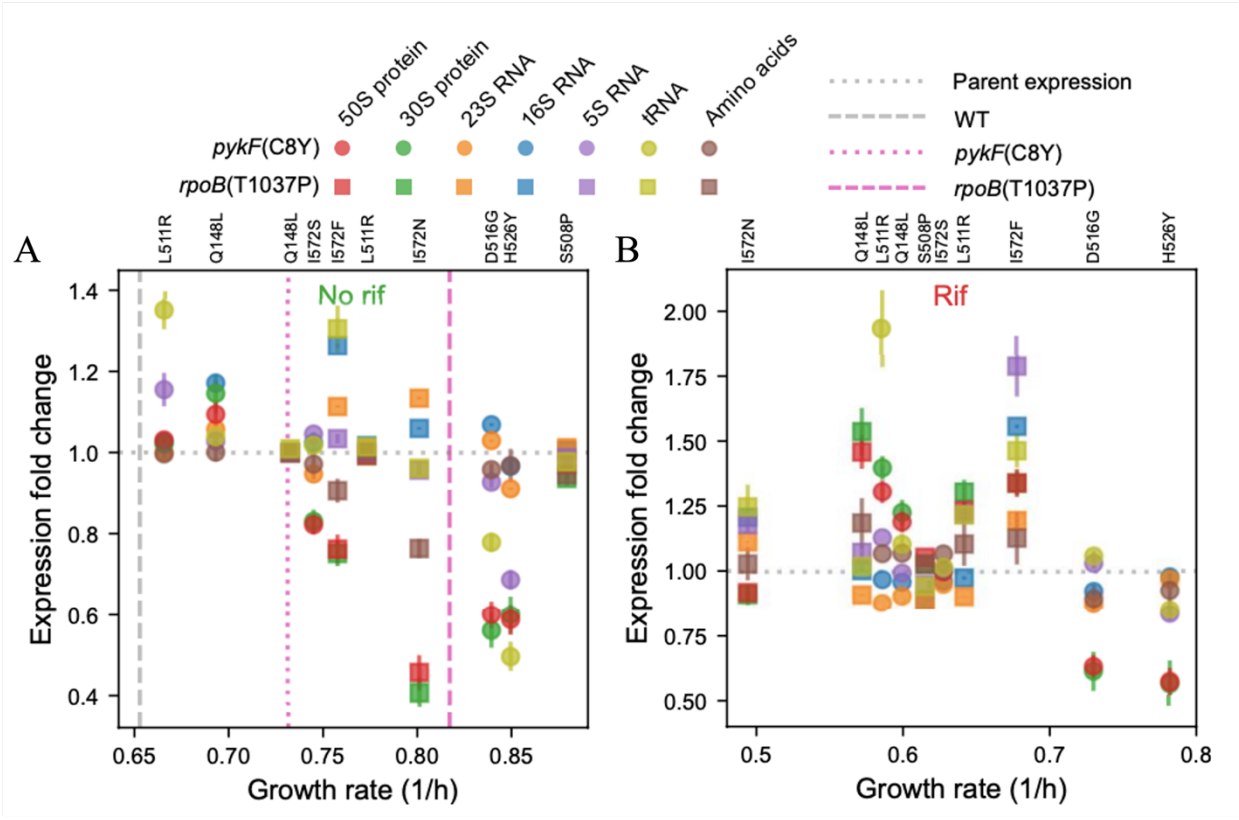

Figure 7: Expression of rRNA loci and transcripts of ribosomal proteins as a function of growth rate.

(A) and (B), Expression fold change for strains cultivated in M9G (A) and M9G + rif (B), where the error bars indicate the standard errors of the mean. The symbols encode the Ref parent strain, and the color indicates the transcript identity for a given Mut strain, as indicated by the corresponding mutation (top text). Wild type and parent growth rates in M9G are indicated by vertical lines and the parental expression is indicated by the horizontal dotted line. The growth rates of sensitive strains are omitted in (B) to show the Mut strain growth rates in more detail.

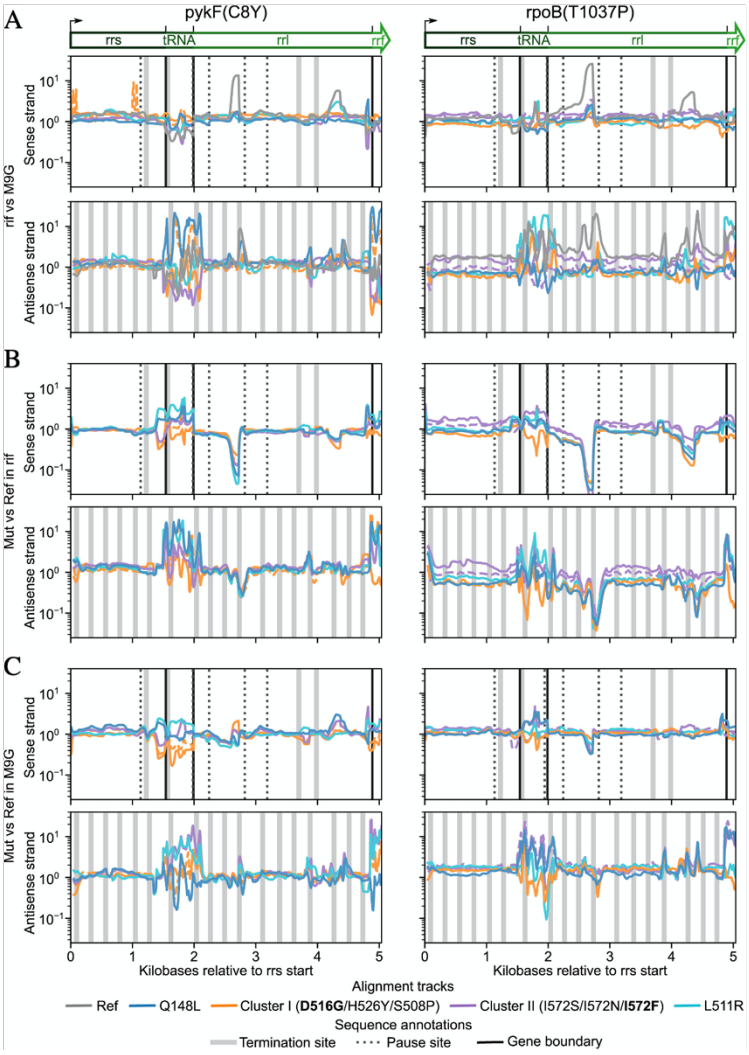

Figure 8: Fold change in RNA-Seq alignments on both strands averaged across all rRNA operons.

(A) Comparison of strains cultivated with and without rif. Data are averaged over the seven rRNA operons whose structure is illustrated by the cartoon above the panel. The top panel shows the sense strand alignment, and the bottom shows the antisense alignment.

(B and C) Same as (A), except comparing Mut strains to their Ref parent when cultivated with rif (B) or without rif (C). Bold face strains have dashed lines in the figures and legends define the meaning of the remaining line colors, color backgrounds, and line styles.

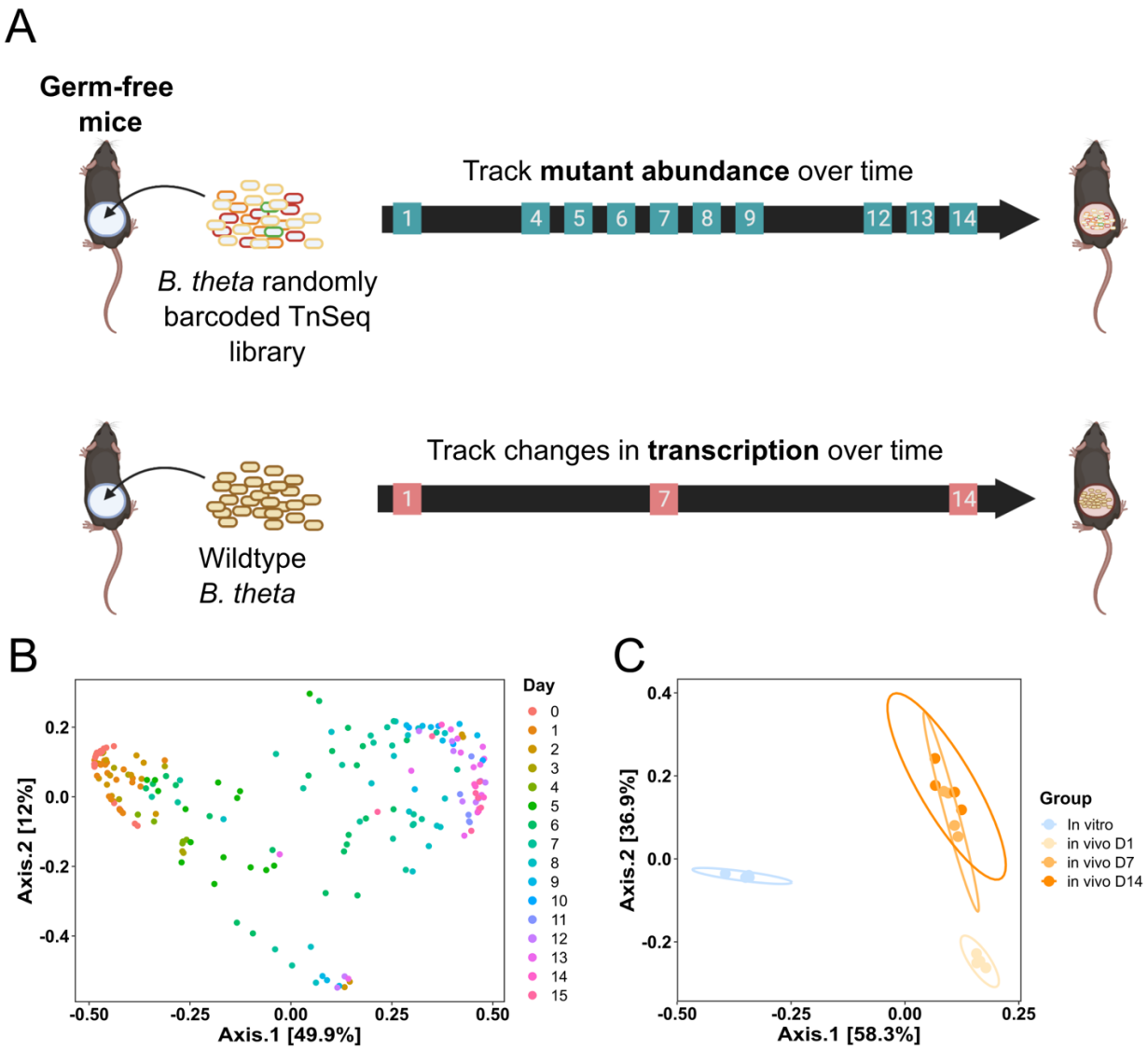

Figure 9: Bt gene expression and genetic fitness determinants shift over the course of colonization and persistence.

A) Experimental scheme of functional genetics (top) and transcriptomics (bottom) experiments in germ-free mice.
B) Principal Coordinates Analysis (PCoA) using Bray-Curtis dissimilarity on the relative abundance of RB-Tn mutant strains within each mouse, colored by experimental day. Four different cohorts of mice (n=3-5 per cohort) are represented. Samples clustered significantly by day ($p = 0.0001$, $R2 = 0.457$, PERMANOVA), experimental run ($p = 0.0001$, $R2 = 0.092$), and mouse ID ($p = 0.0001$, $R2 = 0.135$).
C) PCoA using Bray-Curtis dissimilarity on the expression profiles of Bt within cecal samples at 1, 7, and 14 days after introduction. Bt expression profile from D1 clustered distinctly from that of D7 or D14 (p-value < 0.05, pairwise PERMANOVA), whereas D7 and D14 profiles were statistically indistinguishable.

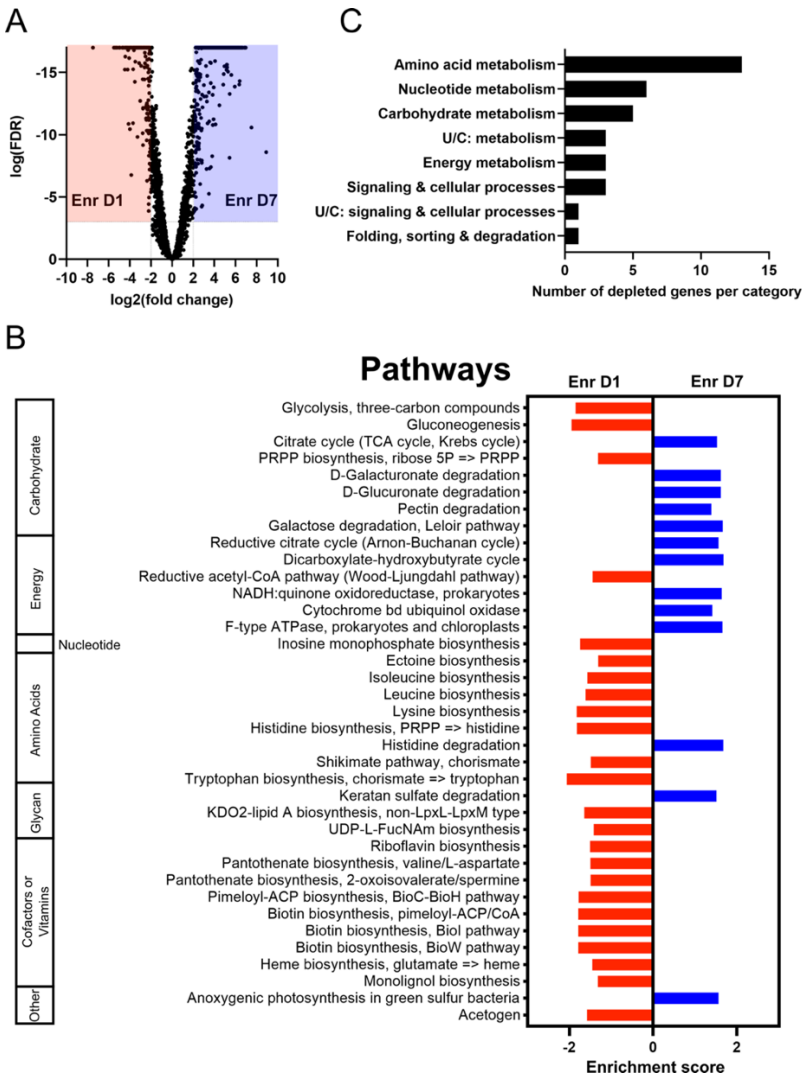

Figure 10: The Bt transcriptome undergoes dramatic remodeling during the first week after introduction to the murine gut.

A) Volcano plot of significant transcriptional differences between D1 and D7 in WT Bt. There were 154 genes significantly enriched on D1 compared to D7 (red), while 359 were significantly enriched on D7 compared to D1 (blue) [cut-off: log(FDR-adjusted p-value) < -3, |log2(fold change)| > 2, and max group mean > 50 TPM].
B) Number of genes mutants significantly depleted in the RB-Tn experiment (t-statistic < -3σ) on D1 of the experiment. Each gene was assigned a KEGG functional category. Only genes with negative fitness are plotted because there were no genes with a positive fitness score / t-statistic on D1.
C) Gene set enrichment analysis (GSEA) of all metabolic pathways in the Kyoto Encyclopedia of Genes and Genomes (KEGG) catalog, comparing D1 to D7 expression. Only pathways with significant differential expression between D1 and D7 ($p<0.05$) are shown.

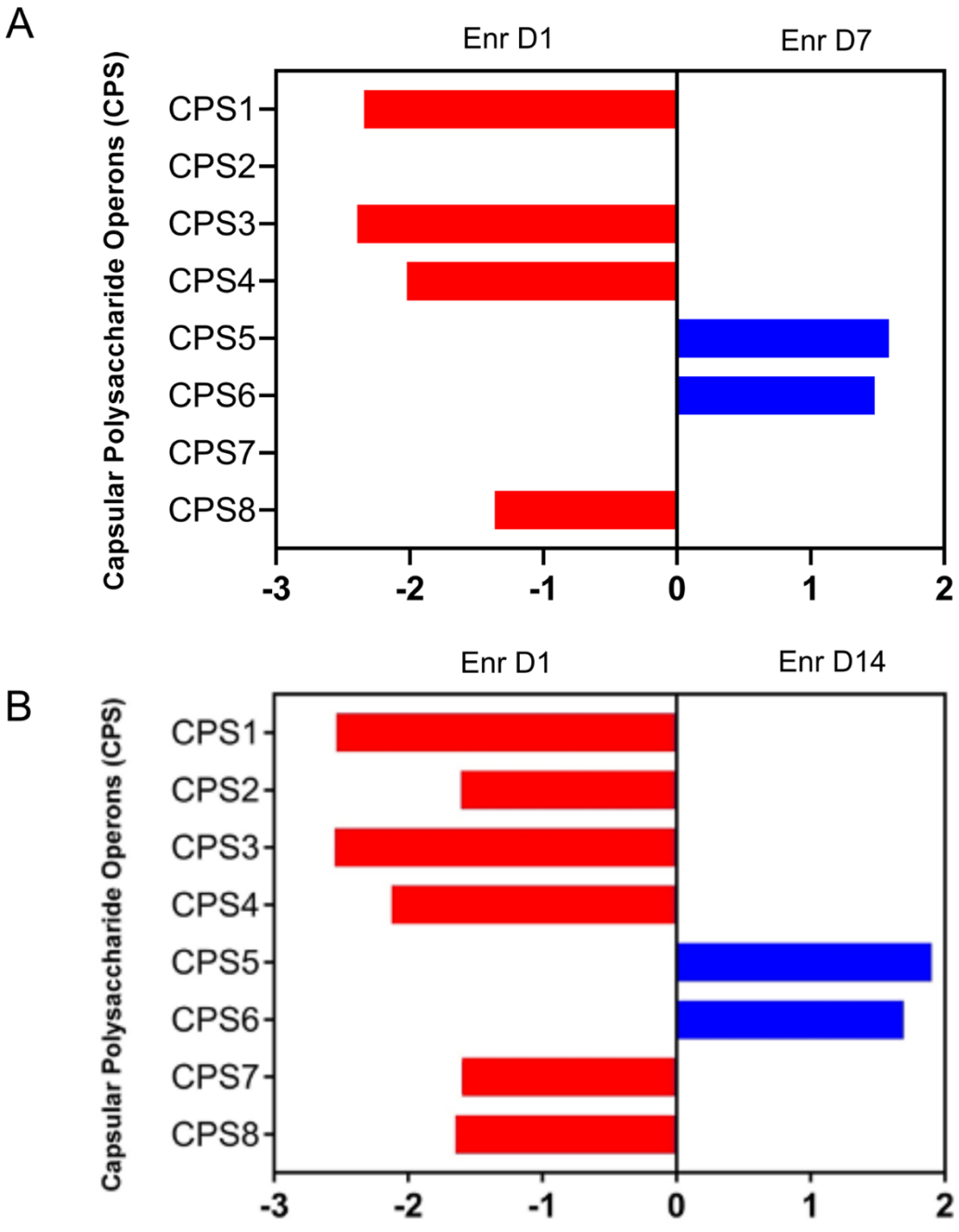

Figure 11: GSEA of capsular polysaccharide biosynthesis operons

GSEA of capsular polysaccharide biosynthesis operons, comparing A) D1 to D7 expression or B) D1 to D14. Only statistically significant scores are shown ($p < 0.05$).

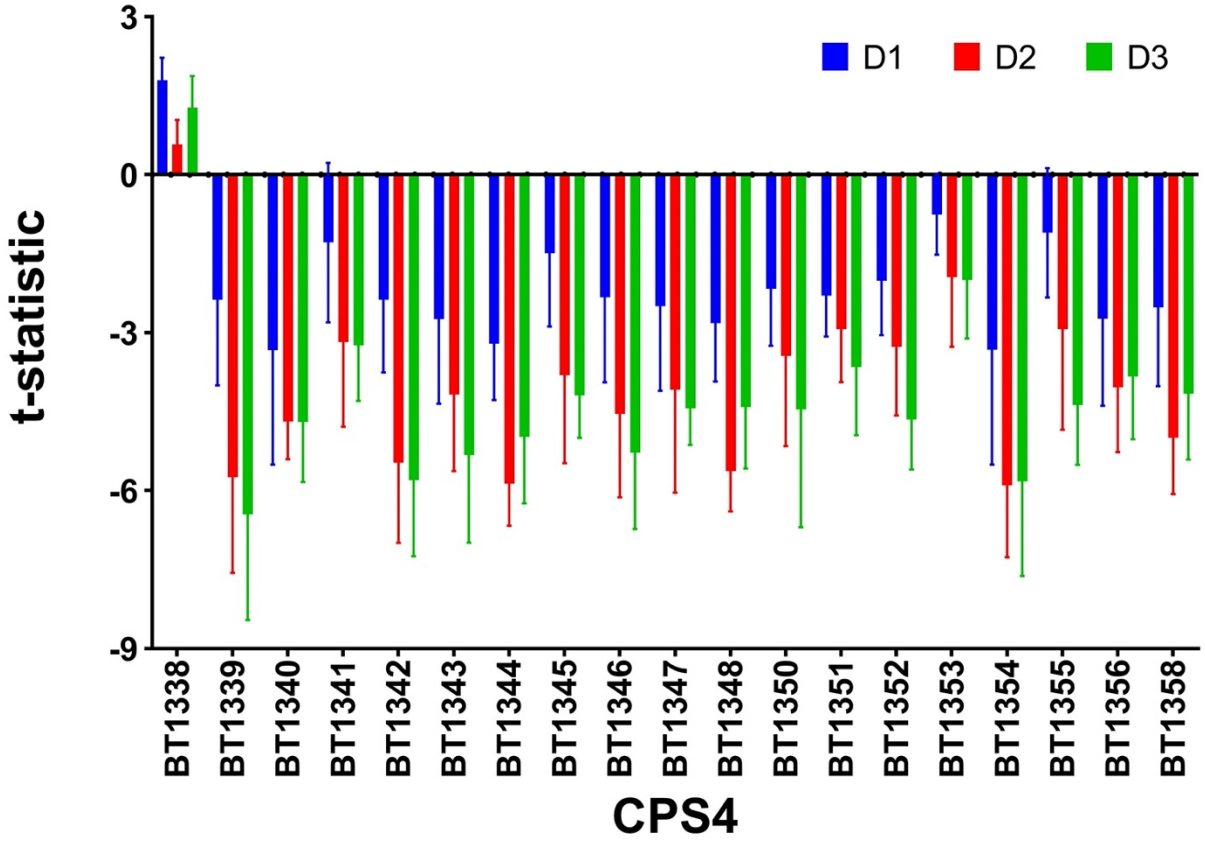

Figure 12: Measurement of the t-statistic for CPS4 (D1-D3)

Measurement of the t-statistic of Wetmore *et al*. in the RB-Tn assay on all days where fitness score was measurable (D1-3) shows that gene insertions in CPS4 consistently results in statistically significant declines in mutant fitness for all genes within the CPS4 locus except for BT1338.

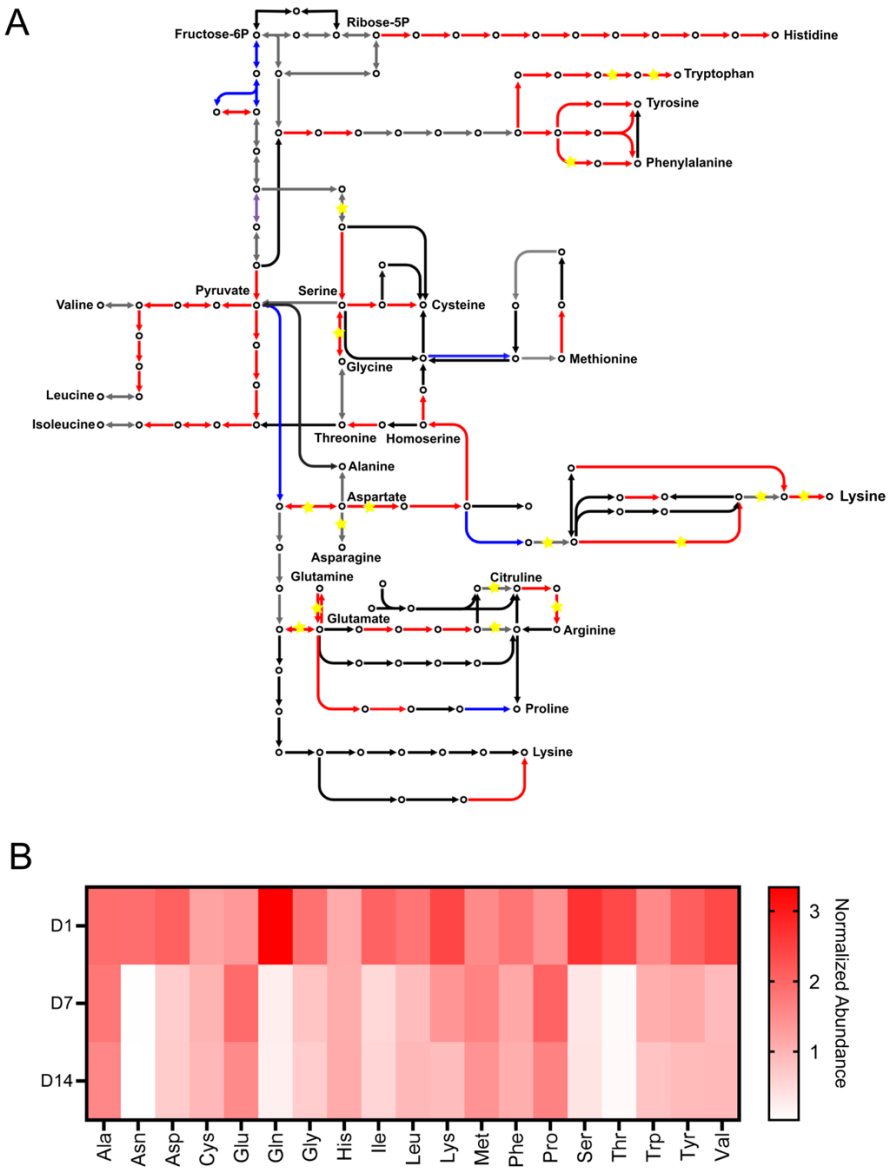

Figure 13: Biosynthesis of amino acids is transcriptionally enriched and functionally significant to early invasion success.

A) Pathway map of reactions related to amino acid biosynthesis. Red arrows represent genes whose transcript is significantly enriched on D1, while blue arrows represent genes whose transcript is significantly enriched on D7 (FDR-adjusted p-value < 0.05). Genes whose transcript level did not differ significantly between D1 and D7 are colored in grey. Black arrows represent reactions for which the associated gene is unknown in the Bt genome. Gold stars represent gene disruptions that were depleted in the RB-Tn assay, as in Fig. 2B.

B) Abundance of various amino acids in the ceca of germ-free mice measured using GCMS and normalized to internal standards and GF standards (n=5 mice for GF, n=4 for post-colonization samples).

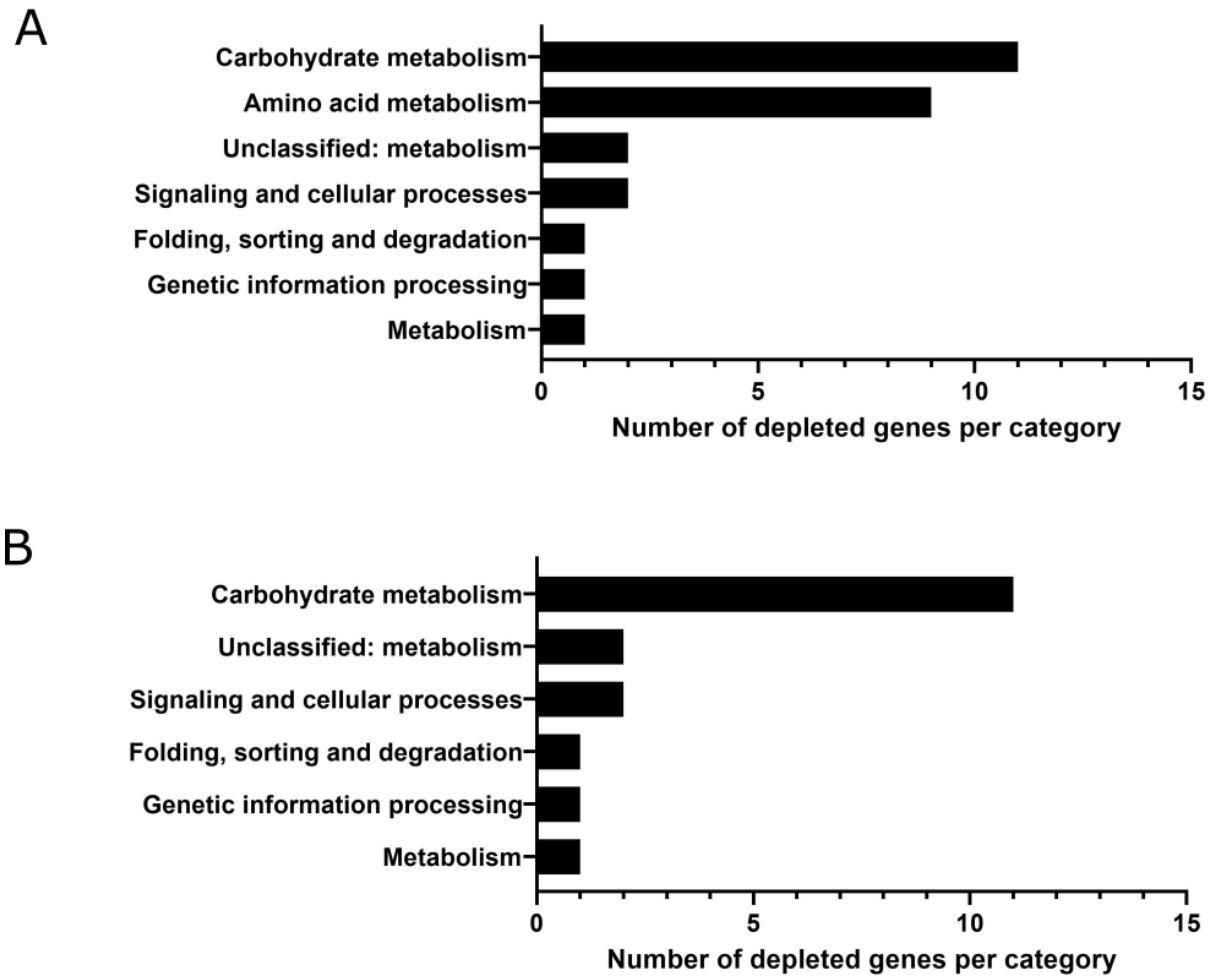

Figure 14: Number of genes mutants significantly depleted early in the RB-Tn experiment.

Each significantly depleted gene (t-statistic < -3σ) on A) D2 or B) D3 was assigned a KEGG functional category if one existed in the database. See Fig. 2D for genes depleted on D1.

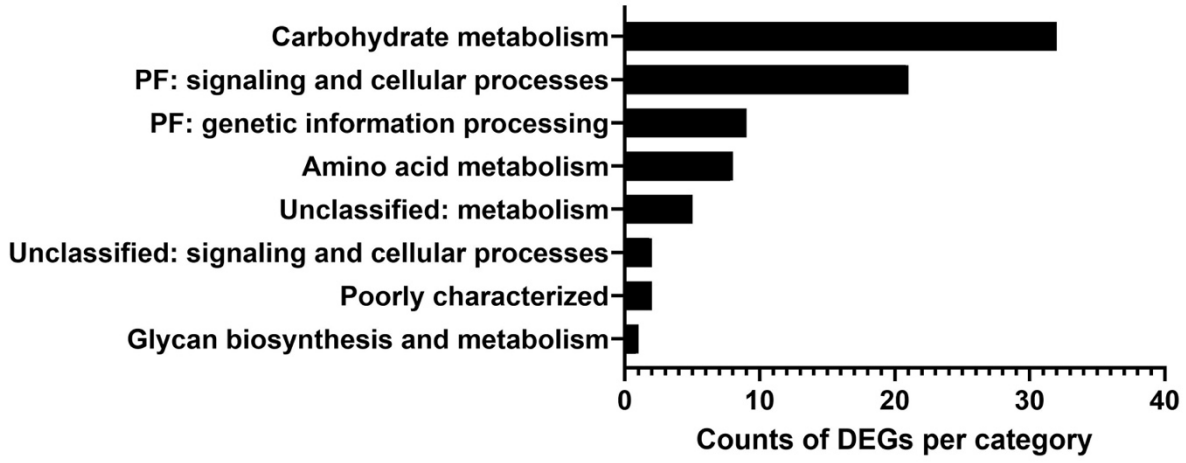

Figure 15: Number of DEGs enriched on D7 compared to D1.

Each gene was assigned a KEGG functional category if one existed in the database.
Cut-off: [log(FDR-adjusted p-value) < -3, |log2(fold change)| > 2, and max group mean > 50 TPM].

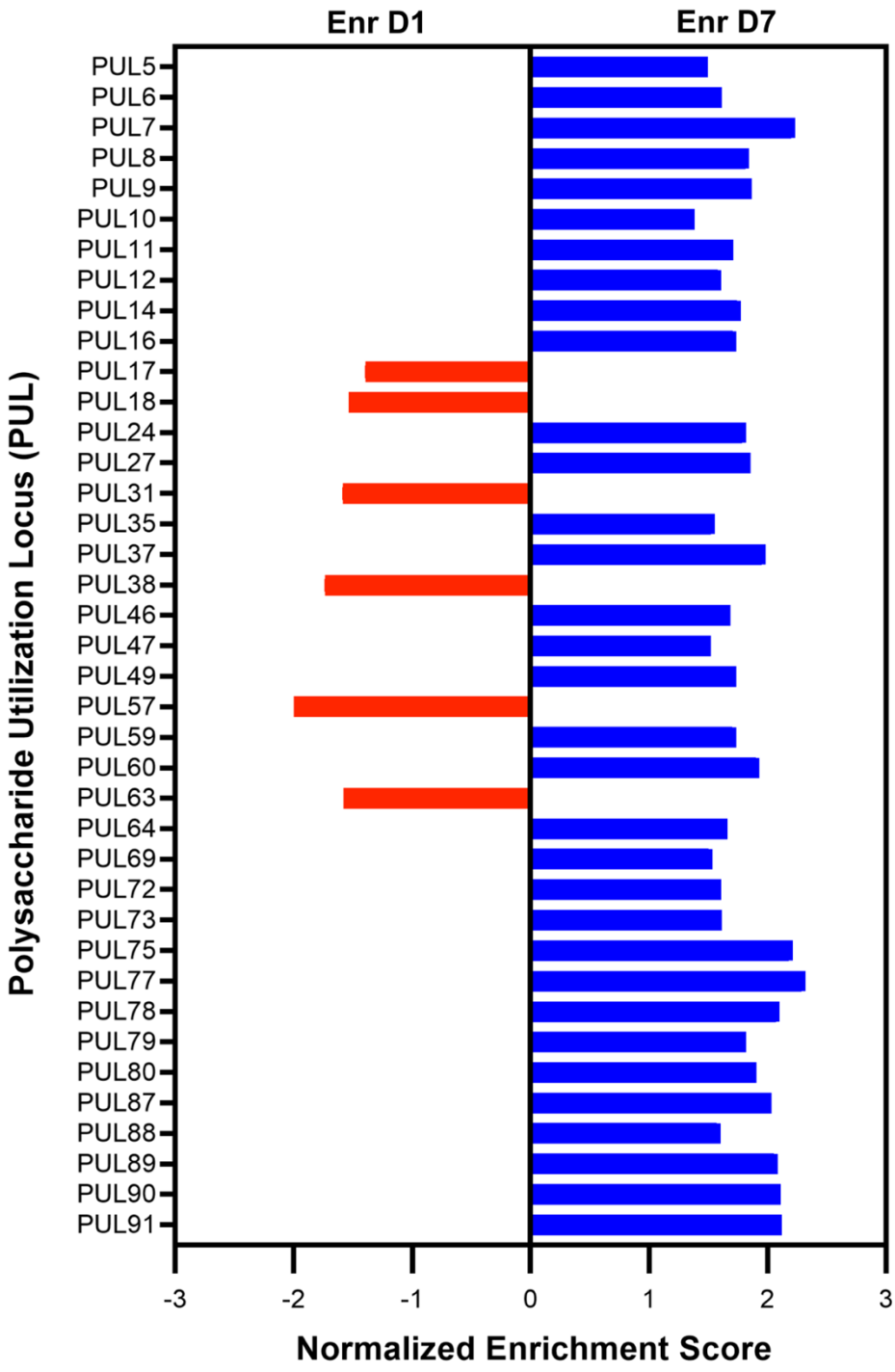

Figure 16: A shift toward higher relative expression of diverse sugar metabolism genes occurs during the first week of gut colonization.

GSEA for genes in the 88 polysaccharide utilization loci of Bt. Only the significant PULs are shown. Of the 42 total PULs that were significantly enriched at any point of the experiment, 34 were enriched on Day 7 or 14.

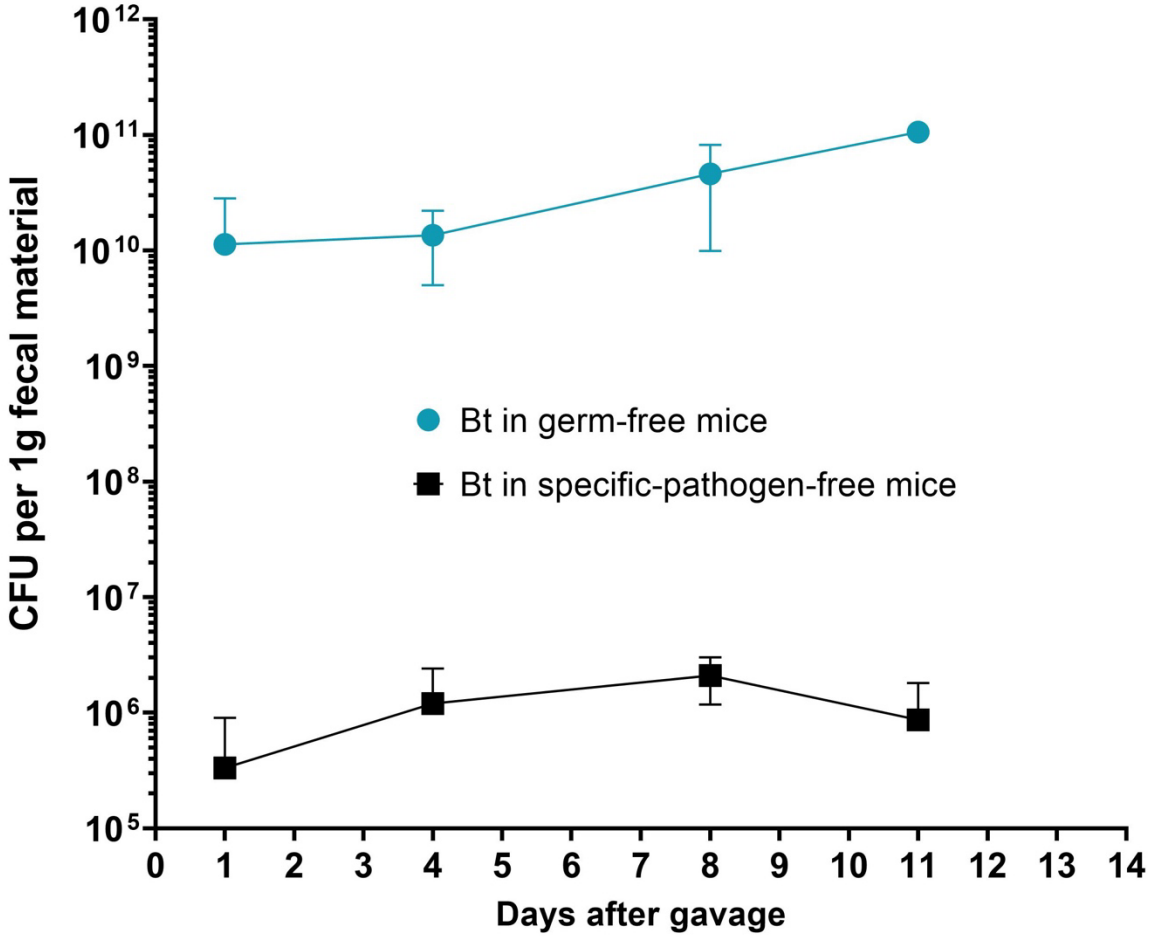

Figure 17: Colonization density for Bt in GF or SPF mice.

Colony forming units (CFU) for RB-Tn mutants in either germ-free (blue) or specific-pathogen-free mice (black). CFU were measured by 1/10 dilution titering on BHI-S plates and counting visible colonies after 48 hours at 37°C in an anaerobic chamber. CFU is normalized for 1g of fecal material.

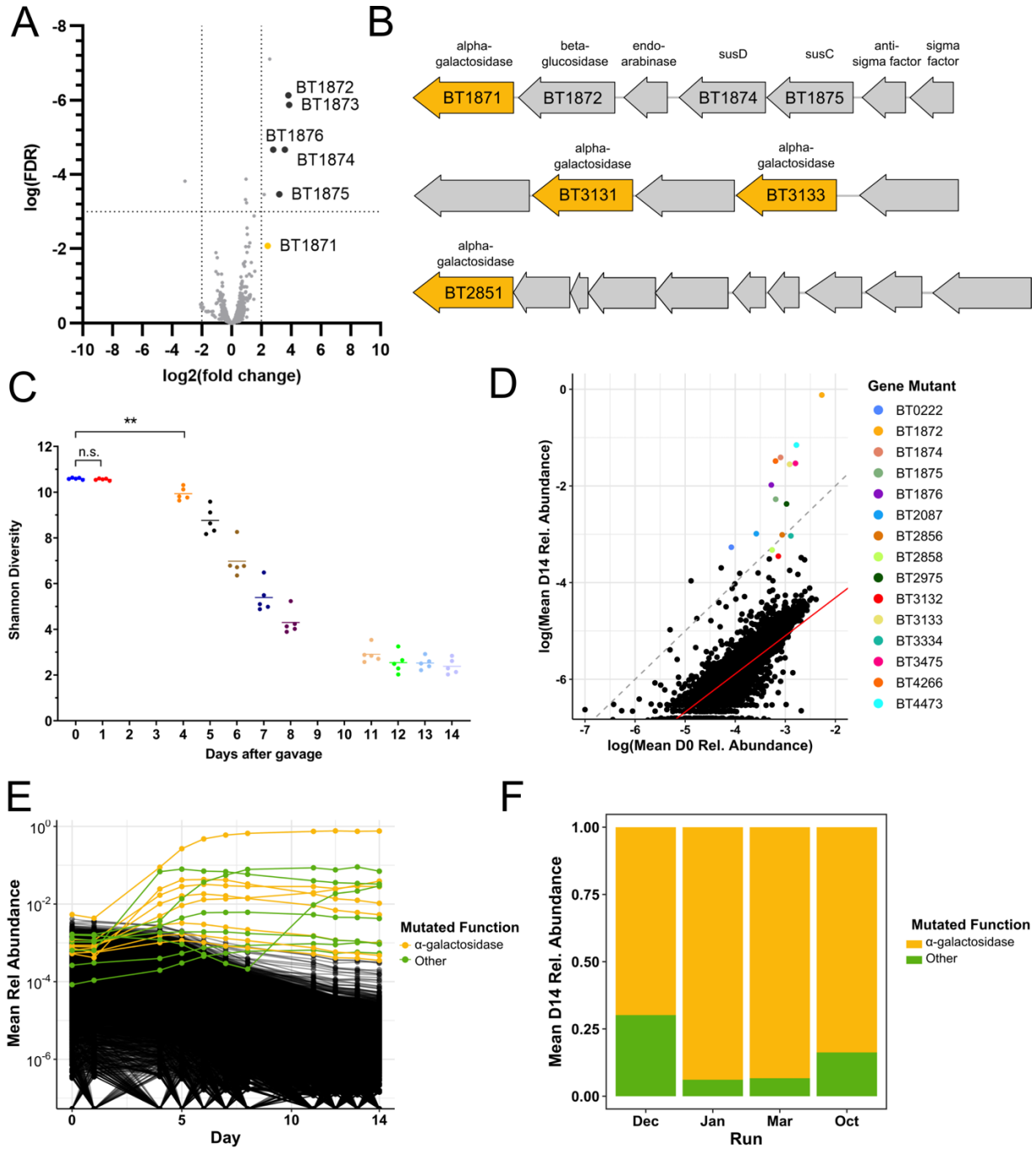

Figure 18: Colonization of a complex Bt mutant library within GF mice selects for disruptions upstream of alpha-galactosidase genes.

A) Volcano plot of significant transcriptional differences between D7 and D14 in WT Bt. Using the same cutoff criteria as in Figure 2A, expression of all genes in PUL24 (red points) except for BT1877 increased significantly throughout the two weeks of the experiment.
B) Organization of PUL 24 (BT1871-1877), PUL 39 (BT2851-2860), and an unnamed PUL containing BT3130-3134. Predicted alpha-galactosidases are colored dark yellow.

Figure 18, continued

C) Shannon diversity of the RB-Tn mutant pool in mice over time. Each point represents the RB-Tn pool within a single mouse on a particular day. The mice are initially colonized by an extremely diverse pool of mutants (300,000+ strains) but diversity drops significantly by D4 (mean = 114,256 strains) and continues to fall throughout the experiment (mean = 20,269 strains by D14).
D) Initial (inoculum) versus final (D14) relative abundance of RB-Tn mutant strains for a single experimental run (October). Each point represents the summed relative abundance of all mutant strains that corresponded to a given gene, averaged across mice. Dashed grey line designates a 1:1 relationship between starting and final abundance; red line represents a linear regression best-fit line generated from the log-transformed data ($p < 2e\text{-}16$, $R2 = 0.6461$, Methods).
E) Relative temporal abundance of RB-Tn mutant strains from a single experimental run (October). Each line represents the summed relative abundance of all mutant strains that mapped to a given gene, averaged across mice. Top 15 most abundant gene mutants at D14 are colored; all others are black.
F) D14 fractional abundance of gene mutants mapping to operons that encode alpha-galactosidase functions (yellow) or other gene functions (green) averaged across mice for each experimental run.

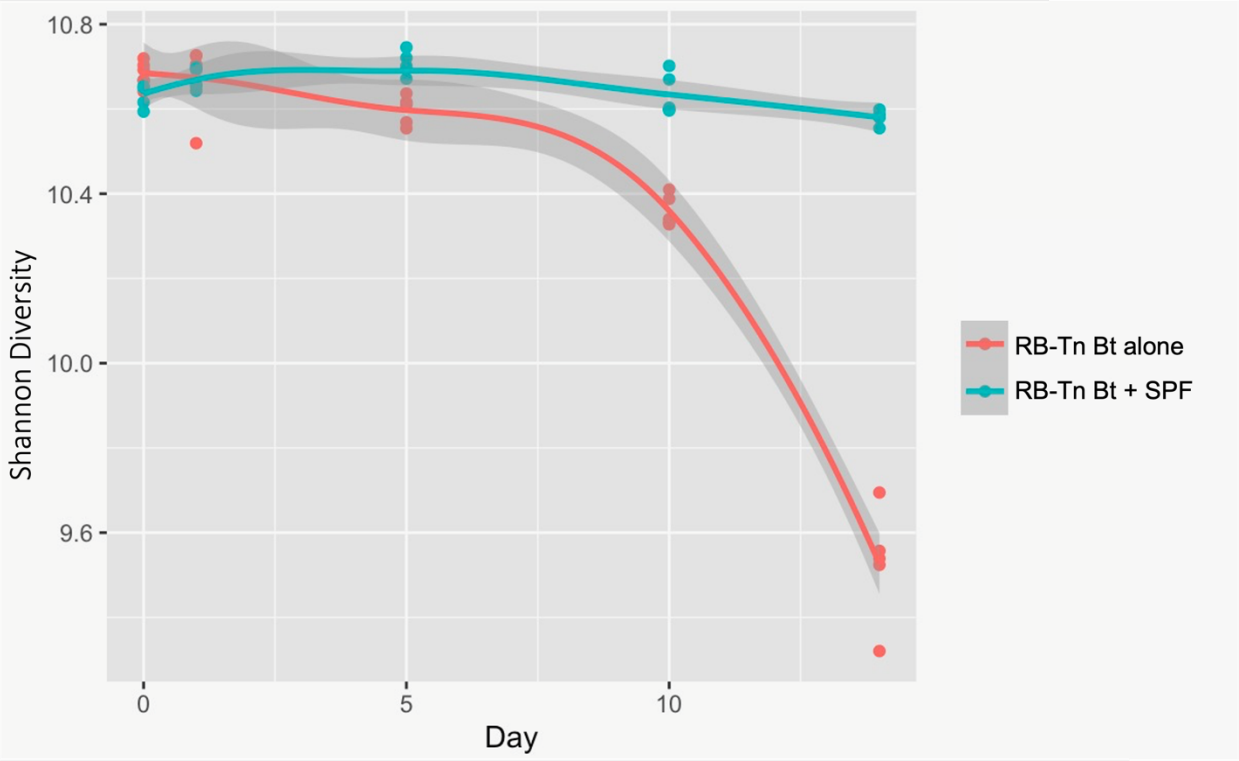

Figure 19: RB-Tn B. theta library remains more diverse when passaged in the presence of complex community than alone.

In BHI-S, Shannon Diversity of a diverse RB-Tn Bt pool decreases significantly after D4 when Bt is alone, but the Bt mutant maintain their diversity over the course of two weeks in the presence of an SPF community. Five cultures of each are maintained for 14 days with ½ dilution every 12 hours.

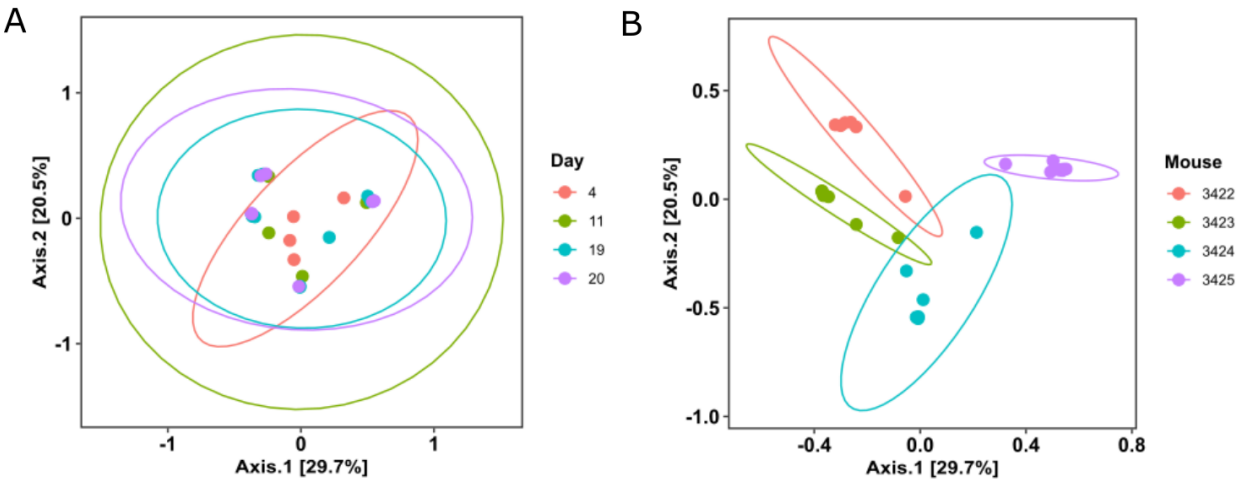

Figure 20: The variance of RB-Tn B. theta mutant population is better accounted for by mouse-to-mouse variation than shifts over time in SPF mice.

Principal Coordinates Analysis (PCoA) using Bray-Curtis dissimilarity on the relative abundance of RB-Tn mutant strains within each mouse, colored by A) experimental day or B) by mouse. Clustering by mouse was significant, whereas clustering by day was not.

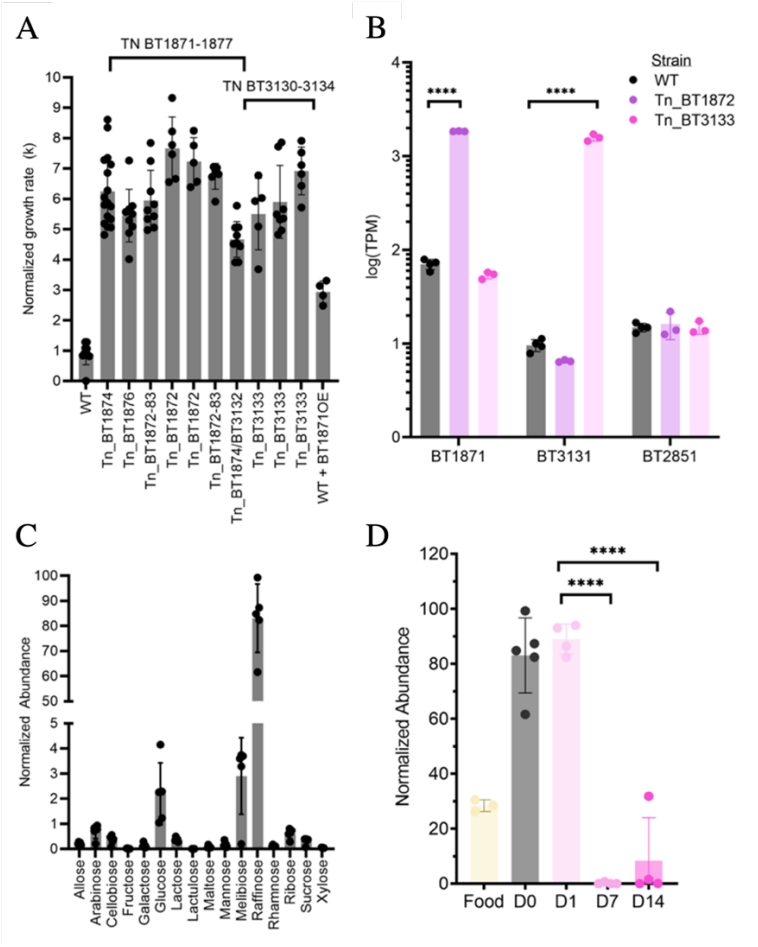

Figure 21: Upregulation of α-galactosidase activity confers a significant growth advantage to Bt in GF mice fed a standard RFO-rich diet.

A) Log phase grow rate (k) of the most abundant RB-Tn mutant strains isolated from mice. Strains with asterisk carry other mutations in addition to the transposon insertion. The WT + BT1871OE strain carries a plasmid copy of BT1871 expressed from the *rpoD* promoter that was integrated into the WT genome at the attN1 site.

B) Normalized abundance of alpha-galactosidase mRNA (transcripts per million, TPM) were measured in WT and hyperfit RB-Tn mutants isolated from mice and grown in Varel-Bryant defined medium with 20mM melibiose as the sole carbon substrate. Expression of the alpha-galactosidase BT1871 was more than 10 times greater with an upstream insertion in BT1872 (Tn_BT1872) compared that of WT, and expression of the alpha-galactosidase BT3131 was more than 100 times greater with an upstream insertion in BT3133 (Tn_BT3133) compared to WT. Expression of alpha-galactosidase BT2851 was unaffected in either of the mutant strains, since the transposon insertions were not upstream of this gene.

C) Abundance of various sugars in the GF mouse ceca as measured using GCMS and normalized to internal standards.

D) Abundance of raffinose in the standard chow fed to GF mice, within GF ceca before colonization, or 1, 7, or 14 days post colonization.

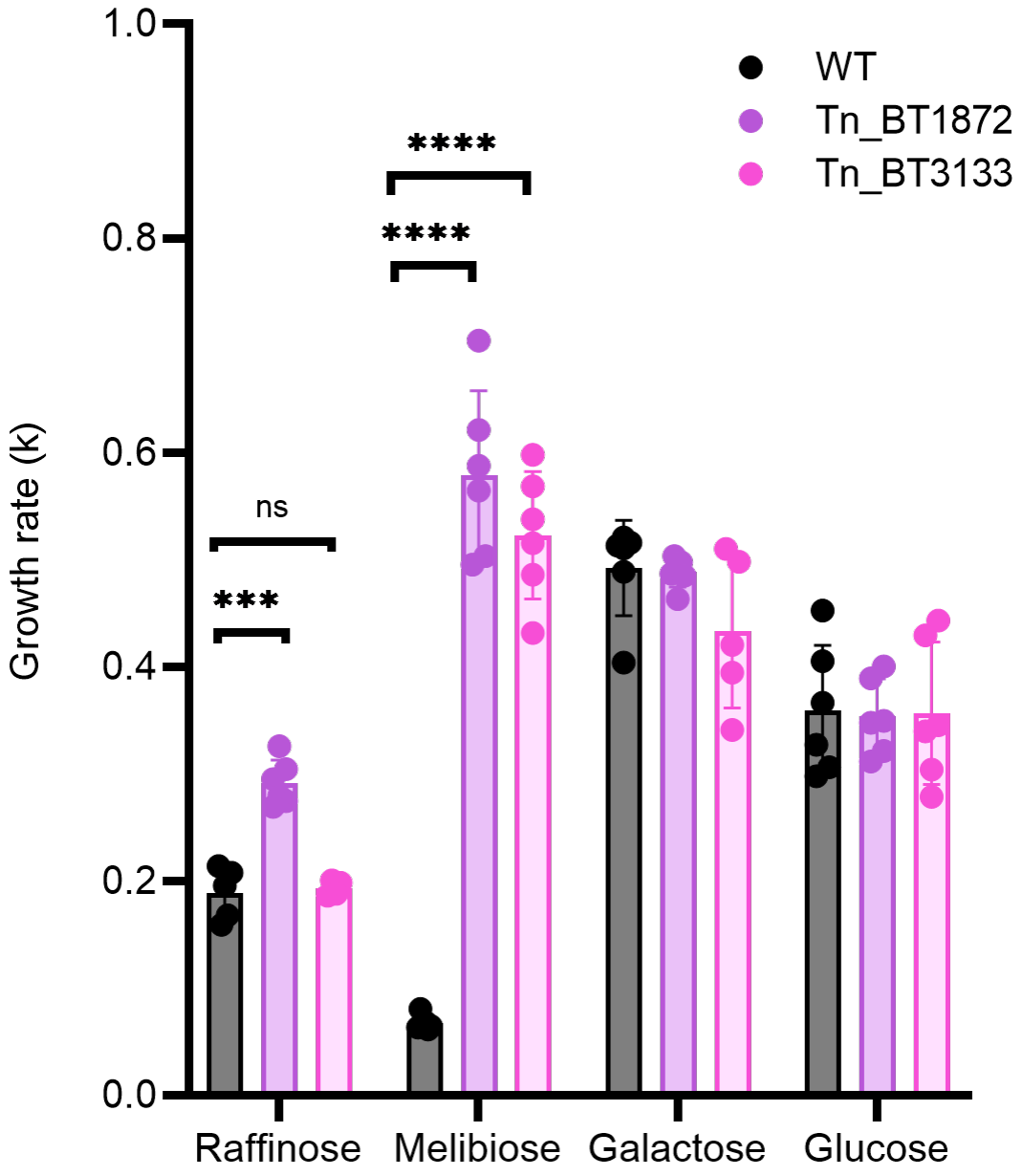

Figure 22: Hyperfit RB-Bt isolates from GF mice grow faster in the presence of 1,6-alpha glycosidic bonded sugars.

The log phase doubling times of Tn_BT3133 and Tn_BT1872 were measured in Varel-Bryant medium with 20mM raffinose, 20mM melibiose, 20mM galactose, or 20mM glucose as the sole carbon substrate.

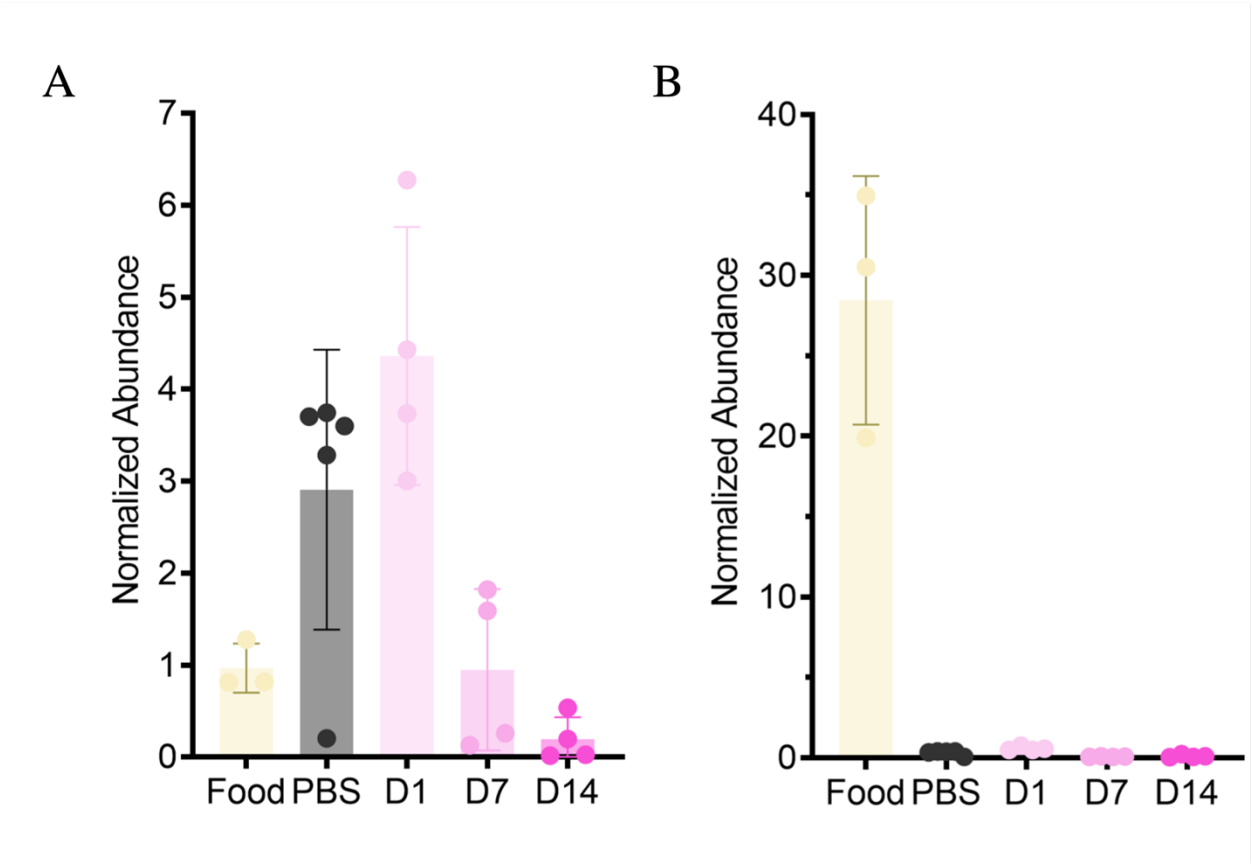

Figure 23: Melibiose is specifically depleted by Bt whereas sucrose is metabolized by the host.

Supplemental Fig. 9: Abundance of A) melibiose or B) sucrose in the standard chow fed to GF mice, within GF ceca before colonization, or 1, 7, or 14 days post colonization.

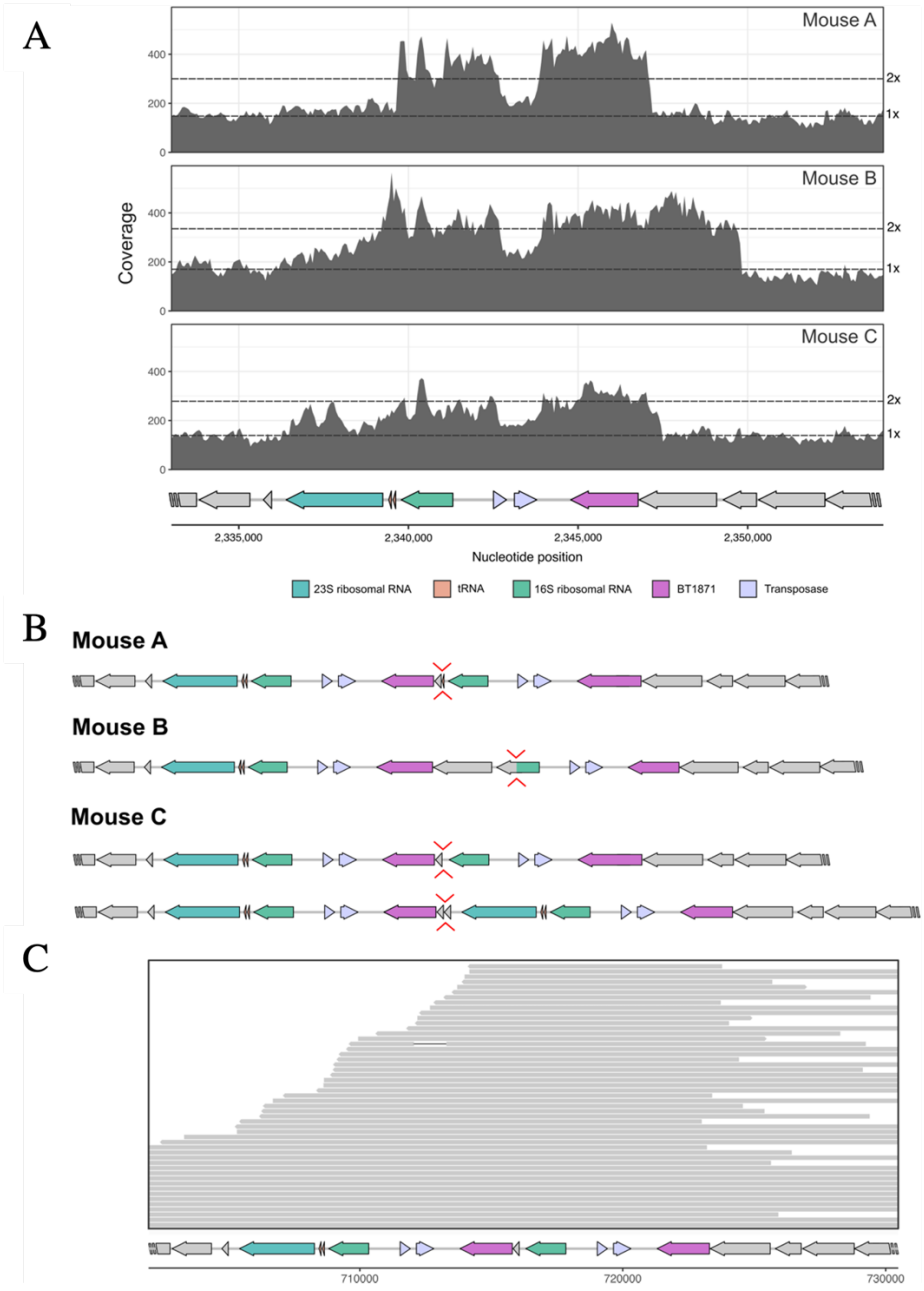

Figure 24: Under strong selective pressures, Bt duplicates the BT1871 locus with the help of a transposable element.

A) Shotgun whole genome sequencing of bulk fecal sample from three mice inoculated with WT Bt for 6 shows at least 2x coverage of the BT1871 locus (RRF = ribosome recycling factor).
B) For each mouse, Breseq analysis of the shotgun whole genome sequencing identified a new junction joining a segment up- and downstream of BT1871 in 60-70% of the reads.
C) Long-read MinION sequencing of an abundant mutant in Mouse C (MZ65) reveals the duplication of a long segment including the BT1871 gene, which generates a tandem repeat of the locus.

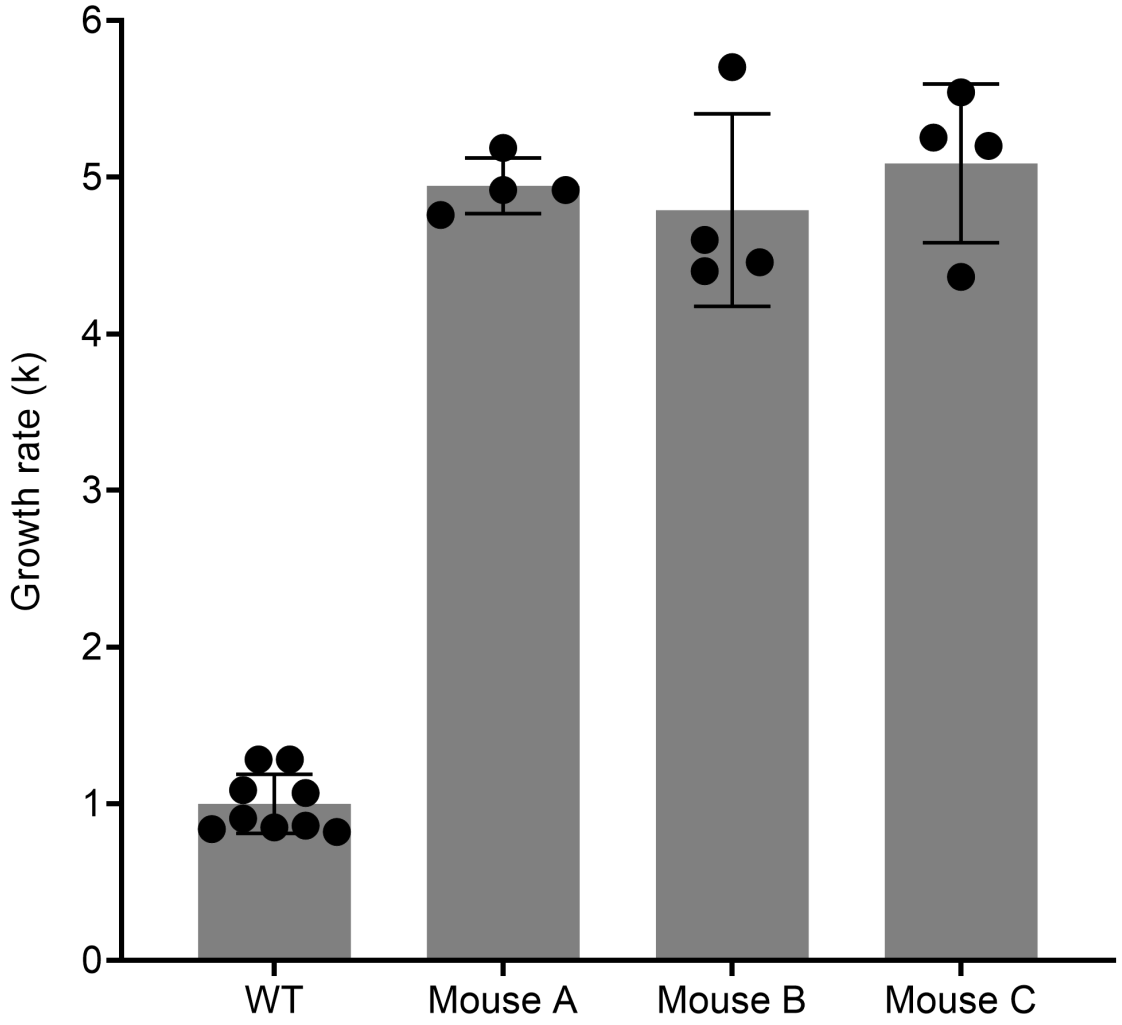

Figure 25: Log phase growth of spontaneous Bt mutants from GF mice is significantly faster than that of WT.

Log phase growth rate (k) of isolated strains from feces of mice colonized with WT Bt for 6 weeks compared to growth rate of WT.

Appendix B – Tables

Table 1: Mutations and Average Doubling Times for *E. coli* Strains Grown on M9G in the Presence and Absence of Rif

Parent Mutation (Ref)	*rpoB* Residue Substitution (Mut)	Rif-resistance Cluster (67)	Doubling Time M9G (min)	Doubling Time M9G + Rif (min)
pykF(C8Y)	–	–	57	120
	Q148L	N	60	70
	L511R	I	63	70
	D516G	I	50	57
	H526Y	I	49	53
	I572S	II	56	67
rpoB(T1037P)	–	–	51	162
	Q148L	N	57	73
	S508P	I	47	68
	L511R	I	54	65
	I572N	II	52	84
	I572F	II	55	62

Table 2: Growth without Rif

M9G Adaptive Mutation	Rif Adaptive Mutation	Rep1	Rep2	Rep3	Rep4
BW25113	None	67	66	60	62
pykF(C8Y)	None	60	57	54	57
pykF(C8Y)	rpoB(H526Y)	48	49	49	50
pykF(C8Y)	rpoB(D516G)	52	50	47	49
pykF(C8Y)	rpoB(L511R)	67	67	57	59
pykF(C8Y)	rpoB(I572S)	55	57	55	55
pykF(C8Y)	rpoB(Q148L)	64	62	55	60
rpoB(T1037P)	None	50	51	51	52
rpoB(T1037P)	rpoB(S508P)	46	46	47	49
rpoB(T1037P)	rpoB(L511R)	51	51	57	57
rpoB(T1037P)	rpoB(I572N)	56	50	50	51
rpoB(T1037P)	rpoB(I572F)	55	55	55	54

Doubling times in minutes for BW25113, pykF(C8Y) and rpoB(T1037P) reference strains, and the Rifampicin-adapted strains grown in M9G.

Table 3: Growth with Rif

M9G Adaptive Mutation	Rif Adaptive Mutation	Rep1	Rep2	Rep3	Rep4
BW25113	None	117	109	120	113
pykF(C8Y)	None	114	138	106	121
pykF(C8Y)	rpoB(H526Y)	53	55	53	51
pykF(C8Y)	rpoB(D516G)	58	55	58	58
pykF(C8Y)	rpoB(L511R)	72	72	70	68
pykF(C8Y)	rpoB(I572S)	68	69	63	68
pykF(C8Y)	rpoB(Q148L)	71	73	67	69
rpoB(T1037P)	None	179	156	152	162
rpoB(T1037P)	rpoB(S508P)	74	67	63	66
rpoB(T1037P)	rpoB(L511R)	65	64	60	71
rpoB(T1037P)	rpoB(I572N)	88	87	81	81
rpoB(T1037P)	rpoB(I572F)	63	58	64	61
rpoB(T1037P)	rpoB(Q148L)	74	78	70	69

Doubling times in minutes for BW25113, pykF(C8Y) and rpoB(T1037P) reference strains, and the Rifampicin-adapted strains grown in M9G with Rif.

Table 4: rpoB(T1037P) analysis

Adaptive mutation	**Growth condition**	**t-test p-value**	**t-statistic**
rpoB(I572F)	M9G + Rif	2.60E-05	-44.04
rpoB(I572N)	M9G + Rif	7.20E-05	-31.20
rpoB(L511R)	M9G + Rif	4.40E-05	-36.89
rpoB(Q148L)	M9G + Rif	8.90E-05	-29.17
rpoB(S508P)	M9G + Rif	1.00E-06	-123.46
rpoB(I572F)	M9G	5.54E-03	7.19
rpoB(I572N)	M9G	5.36E-01	0.70
rpoB(L511R)	M9G	1.48E-01	1.94
rpoB(Q148L)	M9G	2.01E-02	4.53
rpoB(S508P)	M9G	3.58E-03	-8.37

Results of a t-test comparing each growth of each Rifampicin-adapted endpoint with that of the rpoB(T1037P) strain in the presence and absence of Rif. Negative t-values indicate faster growth compared to the parent strain.

Table 5: pykF(C8Y) analysis

Adaptive mutation	**Growth condition**	**t-test p-value**	**t-statistic**
rpoB(D516G)	M9G + Rif	4.95E-04	-16.39
rpoB(H526Y)	M9G + Rif	1.62E-04	-23.81
rpoB(I572S)	M9G + Rif	3.50E-04	-18.41
rpoB(L511R)	M9G + Rif	1.04E-03	-12.74
rpoB(Q148L)	M9G + Rif	4.33E-04	-17.14
rpoB(D516G)	M9G	5.31E-04	-16.00
rpoB(H526Y)	M9G	9.53E-03	-5.94
rpoB(I572S)	M9G	5.50E-01	-0.67
rpoB(L511R)	M9G	4.64E-02	3.28
rpoB(Q148L)	M9G	3.96E-02	3.50

Results of a t-test comparing each growth of each Rifampicin-adapted endpoint with that of the pykF(C8Y) strain in the presence and absence of Rif. Negative t-values indicate faster growth compared to the parent strain.

Table 6: Correlations between transcriptional responses to mutations

	pykF(C8Y)_Q148L_Rif	pykF(C8Y)_L511R_Rif	pykF(C8Y)_I572S_Rif	pykF(C8Y)_D516G_Rif	pykF(C8Y)_H526Y_Rif	rpoB(T1037P)_Q148L_Rif	rpoB(T1037P)_I572F_Rif	rpoB(T1037P)_L511R_Rif	rpoB(T1037P)_I572N_Rif	rpoB(T1037P)_S508P_Rif	pykF(C8Y)_Q148L_M9G	pykF(C8Y)_L511R_M9G	pykF(C8Y)_I572S_M9G	pykF(C8Y)_D516G_M9G	pykF(C8Y)_H526Y_M9G	rpoB(T1037P)_Q148L_M9G	rpoB(T1037P)_I572F_M9G	rpoB(T1037P)_L511R_M9G	rpoB(T1037P)_I572N_M9G	rpoB(T1037P)_S508P_M9G
pykF(C8Y)_Q148L_Rif	1.00	0.49	0.19	-0.23	-0.32	0.39	0.21	0.27	0.00	0.05	0.34	0.24	-0.17	-0.37	-0.40	0.19	0.06	0.06	-0.22	-0.23
pykF(C8Y)_L511R_Rif	0.49	1.00	0.17	-0.12	-0.30	0.32	0.37	0.39	0.19	0.07	0.22	0.59	-0.13	-0.45	-0.61	0.09	0.21	0.20	-0.14	-0.24
pykF(C8Y)_I572S_Rif	0.19	0.17	1.00	0.12	0.06	0.08	0.13	0.15	0.22	0.15	-0.01	0.02	0.11	0.02	-0.04	-0.01	0.00	0.01	0.03	0.00
pykF(C8Y)_D516G_Rif	-0.23	-0.12	0.12	1.00	0.79	-0.50	0.01	-0.27	0.29	-0.05	-0.27	-0.16	0.70	0.68	0.44	-0.14	0.07	0.03	0.31	0.39
pykF(C8Y)_H526Y_Rif	-0.32	-0.30	0.06	0.79	1.00	-0.47	-0.10	-0.32	0.25	0.11	-0.29	-0.37	0.43	0.69	0.67	-0.19	-0.06	-0.12	0.28	0.51
rpoB(T1037P)_Q148L_Rif	0.39	0.32	0.08	-0.50	-0.47	1.00	0.60	0.82	0.36	0.54	0.25	0.11	-0.53	-0.49	-0.44	0.24	0.04	0.17	-0.13	-0.34
rpoB(T1037P)_I572F_Rif	0.21	0.37	0.13	0.01	-0.10	0.60	1.00	0.80	0.77	0.54	0.08	0.19	-0.07	-0.19	-0.37	0.07	0.49	0.39	0.25	-0.13
rpoB(T1037P)_L511R_Rif	0.27	0.39	0.15	-0.27	-0.32	0.82	0.80	1.00	0.56	0.63	0.15	0.23	-0.37	-0.31	-0.39	0.13	0.26	0.42	0.14	-0.19
rpoB(T1037P)_I572N_Rif	0.00	0.19	0.22	0.29	0.25	0.36	0.77	0.56	1.00	0.53	-0.08	0.02	0.16	0.10	-0.07	-0.01	0.45	0.23	0.47	0.11
rpoB(T1037P)_S508P_Rif	0.05	0.07	0.15	-0.05	0.11	0.54	0.54	0.63	0.53	1.00	0.05	-0.09	-0.29	-0.01	0.02	-0.01	0.08	0.14	0.18	0.18
pykF(C8Y)_Q148L_M9G	0.34	0.22	-0.01	-0.27	-0.29	0.25	0.08	0.15	-0.08	0.05	1.00	0.20	-0.19	-0.24	-0.27	0.15	0.03	0.09	-0.14	-0.17
pykF(C8Y)_L511R_M9G	0.24	0.59	0.02	-0.16	-0.37	0.11	0.19	0.23	0.02	-0.09	0.20	1.00	0.02	-0.25	-0.43	0.13	0.38	0.27	0.13	-0.09
pykF(C8Y)_I572S_M9G	-0.17	-0.13	0.11	0.70	0.43	-0.53	-0.07	-0.37	0.16	-0.29	-0.19	0.02	1.00	0.50	0.31	-0.07	0.19	0.04	0.26	0.21
pykF(C8Y)_D516G_M9G	-0.37	-0.45	0.02	0.68	0.69	-0.49	-0.19	-0.31	0.10	-0.01	-0.24	-0.25	0.50	1.00	0.75	-0.15	0.09	-0.12	0.47	0.51
pykF(C8Y)_H526Y_M9G	-0.40	-0.61	-0.04	0.44	0.67	-0.44	-0.37	-0.39	-0.07	0.02	-0.27	-0.43	0.31	0.75	1.00	-0.14	-0.22	-0.28	0.27	0.53
rpoB(T1037P)_Q148L_M9G	0.19	0.09	-0.01	-0.14	-0.19	0.24	0.07	0.13	-0.01	-0.01	0.15	0.13	-0.07	-0.15	-0.14	1.00	0.05	0.09	0.01	-0.10
rpoB(T1037P)_I572F_M9G	0.06	0.21	0.00	0.07	-0.06	0.04	0.49	0.26	0.45	0.08	0.03	0.38	0.19	0.09	-0.22	0.05	1.00	0.33	0.68	0.12
rpoB(T1037P)_L511R_M9G	0.06	0.20	0.01	0.03	-0.12	0.17	0.39	0.42	0.23	0.14	0.09	0.27	0.04	-0.12	-0.28	0.09	0.33	1.00	0.24	-0.01
rpoB(T1037P)_I572N_M9G	-0.22	-0.14	0.03	0.31	0.28	-0.13	0.25	0.14	0.47	0.18	-0.14	0.13	0.26	0.47	0.27	0.01	0.68	0.24	1.00	0.45
rpoB(T1037P)_S508P_M9G	-0.23	-0.24	0.00	0.39	0.51	-0.34	-0.13	-0.19	0.11	0.18	-0.17	-0.09	0.21	0.51	0.53	-0.10	0.12	-0.01	0.45	1.00

Fold change in gene expression was assessed between each Mut strain and its Ref parent. The color-coded correlation coefficients between pairs of log2 fold changes are reported in the table. The formula for row and column identifiers is <parent>_<adaptive mutation in rpoB>_<growth condition>. Strain names are color coded by their growth in M9G relative to the parent strain using the same scheme defined in Fig. 1. The information in the upper triangle of this table is reported in Fig. 3.

Table 7: Strains used in this work

Identifier	Genotype	Source
AMD595	*Bacteroides thetaiotaomicron* VPI-5482 wildtype	Gift from Deutschbauer lab
AMD790	AMD595-derived RB-Tn library	Gift from Deutschbauer lab
Tn_BT1874	RB-Tn mutant isolated from monocolonized mice with barcoded transposon insertion in BT1874	This work
Tn_BT1876	RB-Tn mutant isolated from monocolonized mice with barcoded transposon insertion in BT1876	This work
Tn_BT1872-83	RB-Tn mutant isolated from monocolonized mice with barcoded transposon insertion in the intergenic region between BT1872 and BT1873	This work
Tn_BT1872	RB-Tn mutant isolated from monocolonized mice with barcoded transposon insertion in BT1872	This work
Tn_BT1874/BT3132	RB-Tn mutant isolated from monocolonized mice with barcoded transposon insertion in BT1874 and BT3132	This work
Tn_BT3133	RB-Tn mutant isolated from monocolonized mice with barcoded transposon insertion in BT3133	This work
WT+BT1871OE	WT Bt with a copy of BT1871 integrated into the Bt genome at the attN1 site under the BT1311 (rpoD) constitutive promoter.	This work
MZ65	Bt mutant isolated from feces after 6-week monoassociation of Bt; carries a tandem repeat of the BT1871 locus (see Fig. 7).	This work

Table 8: Custom Ribo-Zero Plus probes

Pool name	Sequence
Supplemental Probe Pool	ATCAACGTCATCGTCTTTAACGACCCTAAGAAATCTAATCTTGTGG CTGG
Supplemental Probe Pool	ACAGTGTCCTCGCAACTACGAGTTAGAACTCAAATAATCAAAGGGC CGTA
Supplemental Probe Pool	CACAAATACTGGCGTACCTGCTTCAAAGTCTCCGGCCTATCCTACA CATC
Supplemental Probe Pool	AGAAAGGAGGTGTTCCAGCCGCACCTTCCGGTACGGCTACCTTGTT ACGA
Supplemental Probe Pool	AGCCAGTATCAACTGCAATTTTACGGTTGAGCCGCAAACTTTCACA ACTG
Supplemental Probe Pool	TATTTGGGGACCTTAGCTGATGGTCTGGATTCTTCTCCTTTAGGACA TGG
Supplemental Probe Pool	CCTCGCATCTTATCAGTCGCTCTACCTCACATTAGTAACTCACAAGG CTG
Supplemental Probe Pool	GCACGGAGTTAGCCGATCCTTATTCATATGGTACATACAAAATTCC ACAC
Supplemental Probe Pool	CGCTCGTTATGGCACTTAAGCCGACACCTCACGGCACGAGCTGACG ACAA
Supplemental Probe Pool	TTGTTAGTCTCCCAGTCAAGCGCCCTTATGCCATTACACTCTACGGA CGG
Supplemental Probe Pool	TACGCTCCCTTTAAACCCAATAAATCCGGATAACGCTCGGATCCTC CGTA
Supplemental Probe Pool	CCAGCGGTCAGTCCAACACGGTCCTCTCGTACTAGTGTCAGAGCCA CGCA
Supplemental Probe Pool	ATCTTACGACGGCAGTCTCTCTAGAGTCCTCAGCATGACCTGTTAGT AAC
Supplemental Probe Pool	GCTGCACATCTCAAGATGCTTAACCTTGCCGGAAAAAGTAACTCGT AGGT
Supplemental Probe Pool	ATACCACTACGCTTAATTTGACTTCGTATACCGGACTATCACCGTCT ATG
Supplemental Probe Pool	AGGGCACCTTTAGAAGCCTCCGTTACTATTTTGGAGGCGACCACCC CAGT
Supplemental Probe Pool	CCTTTATCGTTACTTATACCTACATTTGCTTTTCCACACGCTCCAGC AAA
Supplemental Probe Pool	CGACTTGCATGTGTTAAGCCTGTAGCTAGCGTTCATCCTGAGCCAG GATC
Supplemental Probe Pool	ACGTGTTACTCACCCGTGCGCCGGTCGCCATCTCCAGTTTGCAAGC AAAC
Supplemental Probe Pool	GGATTCTACTCATCAAATGTCTTGCGACATCGTGGTCCTACAACCCC ACA
Supplemental Probe Pool	CCCATATAAAAGAAGTTTACAACCCATAGGGCAGTCATCCTTCACG CTAC

Table 8, continued	
Supplemental Probe Pool	CAATGTGGGGGACCTTCCTCTCAGAACCCCTATCCATCGAAGGTTTGGTG
Supplemental Probe Pool	TAAGGATTAAGTTTAGCGGATTTTCTTGGGAGTATGCTTACACGCACTAT
Supplemental Probe Pool	GAACAGCCCAACCCTTGGGACCTTCTCCAGCCCCAGGATGTGACGAGCCG
Supplemental Probe Pool	TATATCGCAAACAGCGAGTATTCATCGTTTACTGTGTGGACTACCAGGGT
Supplemental Probe Pool	ACATTTGCCTTACGGCTATACTGTTTCCAATATATTCAAATGCAATTTAA
Supplemental Probe Pool	GATTATACCATCGGGTATTAATCTTTCTTTCGAAAGGCTATCCCCGAGTT
Supplemental Probe Pool	AAAGAAAGTGAACGGGCAATTAGTAATGCTCGGCTTTGATGTTACCACCT
Supplemental Probe Pool	CTTTTCAACGCTTATCGGTTCGGTCCTCCAGTTAGTGTTACCTAACCTTC
Supplemental Probe Pool	CCCTCACTCCTGTGATAGAACTAATGCGCATTCGGAGTTTATCAAGACTT
Supplemental Probe Pool	CACGCCGTCACTGCTTAAGCAGCTCCGACCGCTTGTAGGCGCACGGTTTC
Supplemental Probe Pool	CTAACCCTGATCCGATTAGCGTTGATCAGGAAACCTTAGTCTTTCGGCGA
Supplemental Probe Pool	TCGACGCAGAGTGGAATGCTCCCCTACCGATCATTACTGATCCCATAGCT
Supplemental Probe Pool	TTGAGCGATGTCCCTTCCATACGGAAACACCGGATCACTATGCTCTAGTT
Supplemental Probe Pool	ATTTAGCCTTACCGGATGGTCCCGGCAGATTCACGCAAGATTTCTCGTGT
Supplemental Probe Pool	ATACCCACACTTTCGAGCATCAGTGTCAGTTGCAGTCCAGTGAGCTGCCT
Supplemental Probe Pool	GTGGTTTGGGCTATTCCCCGTTCGCTCGCCACTACTAGGGGAATCATTAT
Supplemental Probe Pool	CGATAGATAGAGACCGAACTGTCTCACGACGTTCTGAACCCAGCTCGCGT
Supplemental Probe Pool	ATGCCGGGTTGTCCCATTCGGAAATCTCTGGATCAAAGGTTATTTGCACC
Supplemental Probe Pool	GCCATGGCTGATGCGCGATTACTAGCGAATCCAGCTTCACGAAGTCGGGT
Supplemental Probe Pool	CAGTTTTACCCTAGGACGCTCCTTGCGGTTACGTACTTCAGGTACCCCCG
Supplemental Probe Pool	GCAGCTTATCACGTCCTTCATCGCCTCCGAGAGCCAAGGCATCCGCCATG
Supplemental Probe Pool	GTAATCTTTCAATTTGGCACAGCCCTGTGTTTTTGTTAAACAGTTGCCTG

Table 8, continued	
Supplemental Probe Pool	CTTCTATTCGAAGTGCTTTTCACCTTTCCTTCACAGTACTGGTTCGCTAT
Supplemental Probe Pool	AGTTTCACCGTTGCCGGCGTACTCCCCAGGTGGAATACTTAATGCTTTCG
Supplemental Probe Pool	CGGGGAGTACGAGCTATCTCCAAGTTTGATTAGCCTTTCACCCCCACCCT
Supplemental Probe Pool	AGATCACTTGGTTTCGCGTCTACTCCTTCCGACTCGACGCCCTGTTCAGA
Supplemental Probe Pool	ATGTTAAGCTATAGTAAAGGTTCACGGGGTCTTTTCGTCCCATCGCGGGT
Supplemental Probe Pool	CTTTCAGCACTTATCCAATCCCGACTTAGATACCCAGCAATGCACCTGGC
Supplemental Probe Pool	ATGCTTCTCTTGCGATGACATCTCCTCTTAACCTTCCAGCACCGGGCAGG
Supplemental Probe Pool	CAGGTACTAAGATGTTTCAGTTCCCTGCGTTAGCTTCCATCAAAGATGGA
Supplemental Probe Pool	GGCTGCTTCCAAGCCAACATCCTAGCTGTCTTAGCAATCTGACTTCGTTA
Supplemental Probe Pool	ATGCCCGATTATTATCCACGCCAAACTCCTCGACTAGTGAGCTGTTACGC
Supplemental Probe Pool	CTTCAACGGACTATTCCGTCAGTCCGCGGCGCTGTCACTGCTCCGTCTCC
Supplemental Probe Pool	TTCGTGATATCTAAGCATTTCACCGCTACACCACGAATTCCGCCCACCTC
Supplemental Probe Pool	GAATCACTCAAGCGCCTTAGTATATTCAACCCGACTACGTGTGTCCGTTT
Supplemental Probe Pool	GTAGTACAGGAATATTAACCTGTTCTGCCATCGGCCTCACCGTTCGGCTG
Supplemental Probe Pool	AACGCTCCATACTATCAGGTTCGACTCTCATCCCGGATTTGCCTGGGATG
Supplemental Probe Pool	GAACTGAGAGAGGCTTTTGGGATTAGCATCCTGTCACCAGGTAGCTGCCT
Supplemental Probe Pool	AACTGCCTAATGGAACGCATCCCCATCGATAACCGAAATTCTTTAATAAC
Supplemental Probe Pool	ACCCCTCCGTCGATATGAGCTCTTGGGAGGGATCAGCCTGTTATCCCCGG
Supplemental Probe Pool	GTAACACGTGTGTAGCCCCGGACGTAAGGGCCGTGCTGATTTGACGTCAT
Supplemental Probe Pool	CCCAACTCTCGTTGGGACCCTTTATCCCGAAGTTACAGGGTCAATTTGCC
Supplemental Probe Pool	CTGCGCCCATTGACCAATATTCCTCACTGCTGCCTCCCGTAGGAGTTTGG
Supplemental Probe Pool	CGATACTACAATTTCACTGAGCTCACGGTTGAGACAGTGTCCAGATCATT

Table 9: MinION Sequencing Attributes

	AMD595	MZ55	MZ58	MZ65
Library Kit	SQK-RBK004	SQK-RBK004	SQK-RBK004	SQK-RBK004
Flowcell ID	FAQ98839	FAQ98839	FAQ98839	FAQ98839
Pore Type	R9.4.1 (FLO-MIN106)	R9.4.1 (FLO-MIN106)	R9.4.1 (FLO-MIN106)	R9.4.1 (FLO-MIN106)
Start Materials (ng)	10,000	10,000	10,000	10,000
DNA Shearing	22G needle 10X, 250ul	22G needle 10X, 250ul	22G needle 10X, 250ul	22G needle 10X, 250ul
Barcoding Input (ng)	994.5	792.2	494.7	1249.5
Pool Volume (ul)	5	6	10	4
Pool DNA Input (ng)	497.3	475.3	494.7	499.8
Total Pool Volume (ul)	25	25	25	25
Total Reads	~1,510,000			
Total Pass Reads (%)	1,343,227 (88.95)			
Total Pass Reads per Barcode	344,988	298,543	445,074	113,979
Average length (bp)	7,226	8,857	7,882	8,266
Max. length (bp)	108,747	141,745	114,845	109,375
N50 (bp)	12,596	16,820	14,970	15,675
Top 25% Templates (bp)	9,643	11,730	10,532	11,037
Total Called bases	2,493,077,325	2,644,475,924	3,508,117,512	942,158,294
Base Calling	Guppy v5.0.11	Guppy v5.0.11	Guppy v5.0.11	Guppy v5.0.11
MinKNOW version	4.3.4	4.3.4	4.3.4	4.3.4

Table 10: Metabolomics Internal Standards

Panel	ITSD	Vendor	Product Number	Concentration	Mass/Volume for 1L
SCFA	D3-Acetate	CIL	DLM-3126-25	4 mM	340 mg
SCFA	D5-Propionate	CIL	DLM-1601-1	1 mM	101.1 mg
SCFA	D7-Butyrate	CIL	DLM-7616-PK	.5 mM	58.56 mg
SCFA	D9-Valerate	CIL	DLM-572-5	.25 mM	29.60 μL
SCFA	D8-Valine	CIL	DLM-488-0.25	.05 mM	6.26 mg
SCFA	D6-Succinate	CIL	DLM-831-PK	1 mM	124.14 mg
SCFA	D6-Phenol	CIL	DLM-370-PK	.025 mM	2.5 mg
TMS	U13C-Palmitate	CIL	CLM-409-PK	.04 mM	10.9 mg
TMS/SCFA	D7N15-Proline	CIL	DLNM-7562-0.25	.02 mM	2.46 mg
Bile Acids	D4-Cholic Acid	CIL	DLM-2611-0.05	1 ug/mL	1 mL of 1 mg/mL stock
Bile Acids	D4-Deoxycholic Acid	CIL	DLM-2824-0.01	1 ug/mL	1 mL of 1 mg/mL stock
Bile Acids	D4-Glycocholic Acid	CIL	DLM-2742-0.01	1 ug/mL	1 mL of 1 mg/mL stock
Bile Acids	D4-Glycodeoxycholic Acid	CIL	DLM-9554-0.01	.5 ug/mL	.5 mL of 1 mg/mL stock
Bile Acids	D4-Litchocholic Acid	CIL	DLM-9560-0.05	5 ug/mL	5 mL of 1 mg/mL stock

Table 10, continued					
Bile Acids	D4-Taurocholic Acid	CIL	DLM-9572-0.01	1 ug/mL	1 mL of 1 mg/mL stock
Bile Acids	D4-Taurodeoxycholic Acid	CIL	DLM-9568-0.01	.5 ug/mL	.5 mL of 1 mg/mL stock
Bile Acids	D5-Alpha-Muricholic Acid	CIL	DLM-10627-PK	3.5 ug/mL	3.5 mL of 1 mg/mL stock
Indoles	13C11-L-Tryptophan	CIL	CLM-4290-H	500 nM	50 μL of 10 mM stock
Indoles	13C9-L-Phenylalanine	CIL	CLM-2250-H	500 nM	50 μL of 10 mM stock
Indoles	13C9-L-Tyrosine	CIL	CLM-2263-H	500 nM	50 μL of 10 mM stock
Indoles	D4-Serotonin	CIL	DLM-11030-0	500 nM	50 μL of 10 mM stock
Indoles	13C6-5-Hydroxyindole-3-acetic acid	CIL	CLM-9936-0	500 nM	50 μL of 10 mM stock
Indoles	D4-Tryptamine	CIL	DLM-6989-0	500 nM	50 μL of 10 mM stock
Indoles	D3-Melatonin	CIL	DLM-7101-0	500 nM	50 μL of 10 mM stock
Indoles	13C6-Niacin	CIL	CLM-9954-0	2 μM	200 μL of 10 mM stock
Indoles	13C10-L-Kynurenine	CIL	CLM-9884-0	500 nM	50 μL of 10 mM stock
Indoles	D5-Kynurenic Acid	CIL	DLM-7374-0	500 nM	50 μL of 10 mM stock
Indoles	Anthranilic Acid, Ring-13C6	CIL	CLM-701-0	500 nM	50 μL of 10 mM stock

References

1. Dance A. 2020. Studying life at the extremes. Nature 587:165–166.

2. Pennisi E. 2018. Microbes found in one of Earth's most hostile places, giving hope for life on Mars. Science https://doi.org/10.1126/SCIENCE.AAT4341.

3. Ding J, Zhang Y, Wang H, Jian H, Leng H, Xiao X. 2017. Microbial community structure of deep-sea hydrothermal vents on the ultraslow spreading southwest Indian ridge. Frontiers in Microbiology 8:1012.

4. Antoine L, Bahena-Ceron R, Bunwaree HD, Gobry M, Loegler V, Romby P, Marzi S. 2021. RNA Modifications in Pathogenic Bacteria: Impact on Host Adaptation and Virulence. Genes 2021, Vol 12, Page 1125 12:1125.

5. Martínez LC, Vadyvaloo V. 2014. Mechanisms of post-transcriptional gene regulation in bacterial biofilms. Frontiers in Cellular and Infection Microbiology 5:38.

6. Potts AH, Vakulskas CA, Pannuri A, Yakhnin H, Babitzke P, Romeo T. 2017. Global role of the bacterial post-transcriptional regulator CsrA revealed by integrated transcriptomics. Nature Communications 2017 8:1 8:1–15.

7. López-Maury L, Marguerat S, Bähler J. 2008. Tuning gene expression to changing environments: From rapid responses to evolutionary adaptation. Nature Reviews Genetics 9:583–593.

8. Gottesman S. Post-transcriptional Regulation and the Bacterial Response to Stress. The FASEB Journal 31:22.1-22.1.

9. Patel V, Matange N. 2021. Adaptation and compensation in a bacterial gene regulatory network evolving under antibiotic selection. eLife 10.

10. Tollerson R, Ibba M. 2020. Translational regulation of environmental adaptation in bacteria. The Journal of Biological Chemistry 295:10434.

11. Feugeas JP, Tourret J, Launay A, Bouvet O, Hoede C, Denamur E, Tenaillon O. 2016. Links between Transcription, Environmental Adaptation and Gene Variability in Escherichia coli: Correlations between Gene Expression and Gene Variability Reflect Growth Efficiencies. Molecular Biology and Evolution 33:2515–2529.

12. Retchless AC, Lawrence JG. 2012. Ecological Adaptation in Bacteria: Speciation Driven by Codon Selection. Molecular Biology and Evolution 29:3669–3683.

13. Lenski RE. 2017. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. ISME Journal. Nature Publishing Group https://doi.org/10.1038/ismej.2017.69.

14. Remold SK, Lenski RE. 2004. Pervasive joint influence of epistasis and plasticity on mutational effects in Escherichia coli. Nature genetics 36:423–426.

15. Elena SF, Lenski RE. 2003. Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. Nature Reviews Genetics https://doi.org/10.1038/nrg1088.

16. Morris JJ, Lenski RE, Zinser ER. 2012. The black queen hypothesis: Evolution of dependencies through adaptive gene loss. mBio 3.

17. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. 2007. The Human Microbiome Project. Nature 2007 449:7164 449:804–810.

18. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. 2009. Bacterial community variation in human body habitats across space and time. Science (New York, NY) 326:1694–1697.

19. Dunn RR, Amato KR, Archie EA, Arandjelovic M, Crittenden AN, Nichols LM. 2020. The Internal, External and Extended Microbiomes of Hominins. Frontiers in Ecology and Evolution 8:25.

20. Pickard JM, Zeng MY, Caruso R, Núñez G. 2017. Gut Microbiota: Role in Pathogen Colonization, Immune Responses and Inflammatory Disease. Immunological reviews 279:70.

21. Velazquez EM, Nguyen H, Heasley KT, Saechao CH, Gil LM, Rogers AWL, Miller BM, Rolston MR, Lopez CA, Litvak Y, Liou MJ, Faber F, Bronner DN, Tiffany CR, Byndloss MX, Byndloss AJ, Bäumler AJ. 2019. Endogenous Enterobacteriaceae underlie variation in susceptibility to Salmonella infection. Nature microbiology 4:1057–1064.

22. Dunham J, van Driel N, Eggen BJL, Paul C, T'Hart BA, Laman JD, Kap YS. 2017. Roles of the intestinal microbiota in pathogen protection. Clinical & Translational Immunology 6:e128.

23. Panwar RB, Sequeira RP, Clarke TB. 2021. Microbiota-mediated protection against antibiotic-resistant pathogens. Genes & Immunity 2021 22:5 22:255–267.

24. Miyoshi J, Miyoshi S, Delmont TO, Cham C, Lee STM, Sakatani A, Yang K, Shan Y, Kennedy M, Kiefl E, Yousef M, Crosson S, Sogin M, Antonopoulos DA, Eren AM, Leone V, Chang EB. 2021. Early-Life Microbial Restitution Reduces Colitis Risk Promoted by Antibiotic-Induced Gut Dysbiosis in Interleukin 10–/– Mice. Gastroenterology 161:940-952.e15.

25. Burr AHP, Bhattacharjee A, Hand TW. 2020. Nutritional Modulation of the Microbiome and Immune Response. The Journal of Immunology 205:1479–1487.

26. Lazar V, Ditu LM, Pircalabioru GG, Gheorghe I, Curutiu C, Holban AM, Picu A, Petcu L, Chifiriuc MC. 2018. Aspects of gut microbiota and immune system interactions in infectious diseases, immunopathology, and cancer. Frontiers in Immunology 9:1830.

27. Aidy S el, Dinan TG, Cryan JF. 2014. Immune modulation of the brain-gut-microbe axis. Frontiers in Microbiology 5:146.

28. Belkaid Y, Hand TW. 2014. Role of the Microbiota in Immunity and inflammation. Cell 157:121.

29. Wu HJ, Wu E. 2012. The role of gut microbiota in immune homeostasis and autoimmunity. Gut Microbes 3:4.

30. Zheng D, Liwinski T, Elinav E. 2020. Interaction between microbiota and immunity in health and disease. Cell Research 2020 30:6 30:492–506.

31. Hsieh S, Porter NT, Donermeyer DL, Horvath S, Strout G, Saunders BT, Zhang N, Zinselmeyer B, Martens EC, Stappenbeck TS, Allen PM. 2020. Polysaccharide Capsules Equip the Human Symbiont Bacteroides thetaiotaomicron to Modulate Immune Responses to a Dominant Antigen in the Intestine . The Journal of Immunology 204:1035–1046.

32. Frazier K, Leone V. 2017. Keeping Time in a Relay Race for Fat. Cell host & microbe 22:425.

33. Martin AM, Sun EW, Rogers GB, Keating DJ. 2019. The influence of the gut microbiome on host metabolism through the regulation of gut hormone release. Frontiers in Physiology 10:428.

34. Klassen L, Xing X, Tingley JP, Low KE, King ML, Reintjes G, Abbott DW. 2021. Approaches to Investigate Selective Dietary Polysaccharide Utilization by Human Gut Microbiota at a Functional Level. Frontiers in Microbiology 12:308.

35. Cockburn DW, Koropatkin NM. 2016. Polysaccharide Degradation by the Intestinal Microbiota and Its Influence on Human Health and Disease. Journal of molecular biology 428:3230–3252.

36. Ahmadi S, Mainali R, Nagpal R, Sheikh-Zeinoddin M, Soleimanian-Zad S, Wang S, Deep G, Mishra SK, Yadav H. 2017. Dietary Polysaccharides in the Amelioration of Gut Microbiome Dysbiosis and Metabolic Diseases. Obesity & control therapies : open access 4.

37. Song Q, Wang Y, Huang L, Shen M, Yu Y, Yu Q, Chen Y, Xie J. 2021. Review of the relationships among polysaccharides, gut microbiota, and human health. Food Research International 140:109858.

38. Porter NT, Martens EC. 2017. The Critical Roles of Polysaccharides in Gut Microbial Ecology and Physiology. https://doi.org/101146/annurev-micro-102215-095316 71:349–369.

39. Ndeh D, Baslé A, Strahl H, Yates EA, McClurgg UL, Henrissat B, Terrapon N, Cartmell A. 2020. Metabolism of multiple glycosaminoglycans by Bacteroides thetaiotaomicron is orchestrated by a versatile core genetic locus. Nature Communications 2020 11:1 11:1–12.

40. Ravcheev DA, Godzik A, Osterman AL, Rodionov DA. 2013. Polysaccharides utilization in human gut bacterium Bacteroides thetaiotaomicron: Comparative genomics reconstruction of metabolic and regulatory networks. BMC Genomics 14:1–17.

41. Zhu B, Wang X, Li L. 2010. Human gut microbiome: the second genome of human body. Protein & Cell 1:718.

42. Gill SR, Pop M, DeBoy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, Gordon JI, Relman DA, Fraser-Liggett CM, Nelson KE. 2006. Metagenomic Analysis of the Human Distal Gut Microbiome. Science (New York, NY) 312:1355.

43. Quercia S, Candela M, Giuliani C, Turroni S, Luiselli D, Rampelli S, Brigidi P, Franceschi C, Bacalini MG, Garagnani P, Pirazzini C. 2014. From lifetime to evolution: Timescales of human gut microbiota adaptation. Frontiers in Microbiology. Frontiers Media S.A. https://doi.org/10.3389/fmicb.2014.00587.

44. Bäckhed F, Ding H, Wang T, Hooper L v., Gou YK, Nagy A, Semenkovich CF, Gordon JI. 2004. The gut microbiota as an environmental factor that regulates fat storage. Proceedings of the National Academy of Sciences 101:15718–15723.

45. Clemente JC, Ursell LK, Parfrey LW, Knight R. 2012. The impact of the gut microbiota on human health: an integrative view. Cell 148:1258–1270.

46. Xu L, Surathu A, Raplee I, Chockalingam A, Stewart S, Walker L, Sacks L, Patel V, Li Z, Rouse R. 2020. The effect of antibiotics on the gut microbiome: A metagenomics analysis of microbial shift and gut antibiotic resistance in antibiotic treated mice. BMC Genomics 21:1–18.

47. Modi SR, Collins JJ, Relman DA. 2014. Antibiotics and the gut microbiota. The Journal of clinical investigation 124:4212–4218.

48. Shaw LP, Bassam H, Barnes CP, Walker AS, Klein N, Balloux F. 2019. Modelling microbiome recovery after antibiotics using a stability landscape framework. The ISME Journal 2019 13:7 13:1845–1856.

49. Rinninella E, Raoul P, Cintoni M, Franceschi F, Miggiano GAD, Gasbarrini A, Mele MC. 2019. What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. Microorganisms 7:14.

50. Hillman ET, Lu H, Yao T, Nakatsu CH. 2017. Microbial Ecology along the Gastrointestinal Tract. Microbes and Environments 32:300.

51. Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI. 2008. Worlds within worlds: evolution of the vertebrate gut microbiota. Nature Reviews Microbiology 2008 6:10 6:776–788.

52. Singh S, Verma N, Taneja N. 2019. The human gut resistome: Current concepts & future prospects. The Indian Journal of Medical Research 150:345.

53. Lenski RE. 2017. Convergence and Divergence in a Long-Term Experiment with Bacteria*. https://doi.org/101086/691209 190:S57–S68.

54. Oakley BB, Carbonero F, van der Gast CJ, Hawkins RJ, Purdy KJ. 2010. Evolutionary divergence and biogeography of sympatric niche-differentiated bacterial populations. The ISME Journal 2010 4:4 4:488–497.

55. Zhu Q, Mai U, Pfeiffer W, Janssen S, Asnicar F, Sanders JG, Belda-Ferre P, Al-Ghalith GA, Kopylova E, McDonald D, Kosciolek T, Yin JB, Huang S, Salam N, Jiao JY, Wu Z, Xu ZZ, Cantrell K, Yang Y, Sayyari E, Rabiee M, Morton JT, Podell S, Knights D, Li WJ, Huttenhower C, Segata N, Smarr L, Mirarab S, Knight R. 2019. Phylogenomics of 10,575 genomes reveals evolutionary proximity between domains Bacteria and Archaea. Nature Communications 2019 10:1 10:1–14.

56. Kempher ML, Tao X, Song R, Wu B, Stahl DA, Wall JD, Arkin AP, Zhou A, Zhou J. 2020. Effects of genetic and physiological divergence on the evolution of a sulfate-reducing bacterium under conditions of elevated temperature. mBio 11:1–14.

57. Forbes S, Dobson CB, Humphreys GJ, McBain AJ. 2014. Transient and Sustained Bacterial Adaptation following Repeated Sublethal Exposure to Microbicides and a Novel Human Antimicrobial Peptide. Antimicrobial Agents and Chemotherapy 58:5809.

58. Avrani S, Katz S, Hershberg R. 2020. Adaptations Accumulated under Prolonged Resource Exhaustion Are Highly Transient. mSphere 5.

59. Koskella B, Vos M. 2015. Adaptation in Natural Microbial Populations. http://dx.doi.org/101146/annurev-ecolsys-112414-054458 46:503–522.

60. Susan Slechta E, Liu J, Andersson DI, Roth JR. 2002. Evidence That Selected Amplification of a Bacterial lac Frameshift Allele Stimulates Lac+ Reversion (Adaptive Mutation) With or Without General Hypermutability. Genetics 161:945–956.

61. Dunai A, Spohn R, Farkas Z, Lázár V, Györkei, Apjok G, Boross G, Szappanos B, Grézal G, Faragó A, Bodai L, Papp B, Pál C. 2019. Rapid decline of bacterial drug-resistance in an antibiotic-free environment through phenotypic reversion. eLife 8.

62. Cairns J, Foster PL. 1991. Adaptive Reversion of a Frameshift Mutation in Escherichia Coli. Genetics 128:695.

63. Linhart YB, Grant MC. 2003. EVOLUTIONARY SIGNIFICANCE OF LOCAL GENETIC DIFFERENTIATION IN PLANTS. http://dx.doi.org/101146/annurev.ecolsys271237 27:237–277.

64. Hawks J. 2012. Longer time scale for human evolution. Proceedings of the National Academy of Sciences 109:15531–15532.

65. Lieberman TD, Flett KB, Yelin I, Martin TR, McAdam AJ, Priebe GP, Kishony R. 2014. Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressures. Nature genetics 46:82.

66. Lieberman TD, Michel JB, Aingaran M, Potter-Bynoe G, Roux D, Davis MR, Skurnik D, Leiby N, Lipuma JJ, Goldberg JB, McAdam AJ, Priebe GP, Kishony R. 2011. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. Nature genetics 43:1275.

67. Díaz-Raviña M, Bååth E. 1996. Development of metal tolerance in soil bacterial communities exposed to experimentally increased metal levels. Applied and environmental microbiology 62:2970–2977.

68. Vasquez KS, Willis L, Cira NJ, Ng KM, Pedro MF, Aranda-Díaz A, Rajendram M, Yu FB, Higginbottom SK, Neff N, Sherlock G, Xavier KB, Quake SR, Sonnenburg JL, Good BH, Huang KC. 2021. Quantifying rapid bacterial evolution and transmission within the mouse intestine. Cell host & microbe 29:1454-1468.e4.

69. Buckling A, Rainey PB. 2002. Antagonistic coevolution between a bacterium and a bacteriophage. Proceedings of the Royal Society B: Biological Sciences 269:931.

70. Remold SK, Lenski RE. 2004. Pervasive joint influence of epistasis and plasticity on mutational effects in Escherichia coli. Center for Microbial Ecology, Michigan State University, East Lansing, 48824, USA. susanna.remold\char64yale.edu https://doi.org/10.1038/ng1324.

71. Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE. 2006. Tests of parallel molecular evolution in a long-term experiment with Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America 103:9107–9112.

72. Engelhardt D, Shakhnovich EI. 2019. Mutation rate variability as a driving force in adaptive evolution. Physical Review E 99:022424.

73. Eyre-Walker A, Keightley PD. 2007. The distribution of fitness effects of new mutations. Nature reviews Genetics 8:610–618.

74. Oliver A, Cantón R, Campo P, Baquero F, Blázquez J. 2000. High Frequency of Hypermutable Pseudomonas aeruginosa in Cystic Fibrosis Lung Infection. Science 288:1251–1253.

75. Matic I, Radman M, Taddei F, Picard B, Doit C, Bingen E, Denamur E, Elion J, Leclerc JE, Cebula TA. 1997. Highly variable mutation rates in commensal and pathogenic Escherichia coli. Science 277:1833–1834.

76. Harris SR, Feil EJ, Holden MTG, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ, Bentley SD. 2010. Evolution of MRSA during hospital transmission and intercontinental spread. Science (New York, NY) 327:469–474.

77. Smith EE, Buckley DG, Wu Z, Saenphimmachak C, Hoffman LR, D'Argenio DA, Miller SI, Ramsey BW, Speert DP, Moskowitz SM, Burns JL, Kaul R, Olson M v. 2006. Genetic adaptation by Pseudomonas aeruginosa to the airways of cystic fibrosis patients. Proceedings of the National Academy of Sciences of the United States of America 103:8487–8492.

78. Tasnim N, Abulizi N, Pither J, Hart MM, Gibson DL. 2017. Linking the gut microbial ecosystem with the environment: Does gut health depend on where we live? Frontiers in Microbiology 8:1935.

79. Donaldson GP, Lee SM, Mazmanian SK. 2016. Gut biogeography of the bacterial microbiota. Nature reviews Microbiology 14:20.

80. Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine MC, Kane M, Krumins JA, Kuske CR, Morin PJ, Naeem S, Øvreås L, Reysenbach AL, Smith VH, Staley JT. 2006. Microbial biogeography: putting microorganisms on the map. Nature Reviews Microbiology 2006 4:2 4:102–112.

81. Fierer N, Jackson RB. 2006. The diversity and biogeography of soil bacterial communities. Proceedings of the National Academy of Sciences 103:626–631.

82. Stubbendieck RM, Vargas-Bautista C, Straight PD. 2016. Bacterial communities: Interactions to scale. Frontiers in Microbiology 7:1234.

83. Velsko IM, Fellows Yates JA, Aron F, Hagan RW, Frantz LAF, Loe L, Martinez JBR, Chaves E, Gosden C, Larson G, Warinner C. 2019. Microbial differences between dental plaque and historic dental calculus are related to oral biofilm maturation stage. Microbiome 7:1–20.

84. Tropini C, Earle KA, Huang KC, Sonnenburg JL. 2017. The gut microbiome: Connecting spatial organization to function. Cell host & microbe 21:433.

85. Claudino ES, Lyra ML, Gleria I, Campos PRA. 2013. Adaptive evolution on a continuous lattice model. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 87:032711.

86. Perfeito L, Pereira MI, Campos PRA, Gordo I. 2008. The effect of spatial structure on adaptation in Escherichia coli. Biology letters 4:57–59.

87. Habets MGJL, Czárán T, Hoekstra RF, de Visser JAGM. 2007. Spatial structure inhibits the rate of invasion of beneficial mutations in asexual populations. Proceedings Biological sciences 274:2139–2143.

88. Gordo I, Campos PRA. 2006. Adaptive evolution in a spatially structured asexual population. Genetica 127:217–229.

89. Collins S, de Meaux J, Acquisti C. 2007. Adaptive Walks Toward a Moving Optimum. Genetics 176:1089–1099.

90. Pease CM, Lande R, Bull JJ. 1989. A Model of Population Growth, Dispersal and Evolution in a Changing Environment. Ecology 70:1657–1664.

91. Guzella TS, Dey S, Chelo IM, Pino-Querido A, Pereira VF, Proulx SR, Teotónio H. 2018. Slower environmental change hinders adaptation from standing genetic variation. PLOS Genetics 14:e1007731.

92. Perron GG, Gonzalez A, Buckling A. 2008. The rate of environmental change drives adaptation to an antibiotic sink. Journal of Evolutionary Biology 21:1724–1731.

93. Darmon E, Leach DRF. 2014. Bacterial Genome Instability. Microbiology and Molecular Biology Reviews 78:1–39.

94. Denamur E, Matic I. 2006. Evolution of mutation rates in bacteria. Molecular Microbiology 60:820–827.

95. Vasu K, Nagaraja V. 2013. Diverse Functions of Restriction-Modification Systems in Addition to Cellular Defense. Microbiology and Molecular Biology Reviews : MMBR 77:53.

96. Hacker J, Carniel E. 2001. Ecological fitness, genomic islands and bacterial pathogenicity. EMBO reports 2:376–381.

97. Litsios A, Ortega ÁD, Wit EC, Heinemann M. 2018. Metabolic-flux dependent regulation of microbial physiology. Current Opinion in Microbiology 42:71–78.

98. Béchon N, Mihajlovic J, Vendrell-Fernández S, Chain F, Langella P, Beloin C, Ghigo JM. 2020. Capsular polysaccharide cross-regulation modulates bacteroides thetaiotaomicron biofilm formation. mBio 11:1–14.

99. Porter NT, Canales P, Peterson DA, Martens EC. 2017. A subset of polysaccharide capsules in the human symbiont Bacteroides thetaiotaomicron promote increased competitive fitness in the mouse gut. Cell host & microbe 22:494.

100. Fiebig A, Herrou J, Fumeaux C, Radhakrishnan SK, Viollier PH, Crosson S. 2014. A Cell Cycle and Nutritional Checkpoint Controlling Bacterial Surface Adhesion. PLOS Genetics 10:e1004101.

101. Patel S. 2016. Drivers of bacterial genomes plasticity and roles they play in pathogen virulence, persistence and drug resistance. Infection, Genetics and Evolution. Elsevier B.V. https://doi.org/10.1016/j.meegid.2016.08.030.

102. Sabrina Pankey M, Foxall RL, Ster IM, Perry LA, Schuster BM, Donner RA, Coyle M, Cooper VS, Whistler CA. 2017. Host-selected mutations converging on a global regulator drive an adaptive leap towards symbiosis in bacteria. eLife 6.

103. Maunders EA, Triniman RC, Western J, Rahman T, Welch M. 2020. Global reprogramming of virulence and antibiotic resistance in Pseudomonas aeruginosa by a single nucleotide polymorphism in elongation factor, fusA1. Journal of Biological Chemistry 295:16411–16426.

104. Ali F, Seshasayee ASN. 2020. Dynamics of genetic variation in transcription factors and its implications for the evolution of regulatory networks in Bacteria. Nucleic Acids Research 48:4100.

105. Elena SF, Lenski RE. 1997. Test of synergistic interactions among deleterious mutations in bacteria. Nature 390:395–398.

106. Turkarslan S, Stopnisek N, Thompson AW, Arens CE, Valenzuela JJ, Wilson J, Hunt KA, Hardwicke J, de Lomana ALG, Lim S, Seah YM, Fu Y, Wu L, Zhou J, Hillesland KL, Stahl DA, Baliga NS. 2021. Synergistic epistasis enhances the co-operativity of mutualistic interspecies interactions. The ISME Journal 2021 15:8 15:2233–2247.

107. Apidianakis Y, Pitsouli C, Perrimon N, Rahme L. 2009. Synergy between bacterial infection and genetic predisposition in intestinal dysplasia. Proceedings of the National Academy of Sciences of the United States of America 106:20883–20888.

108. Duong L, Gross SP, Siryaporn A. 2021. Developing Antimicrobial Synergy with AMPs. Frontiers in Medical Technology 0:9.

109. Lalejini A, Ferguson AJ, Grant NA, Ofria C. 2021. Adaptive Phenotypic Plasticity Stabilizes Evolution in Fluctuating Environments. Frontiers in Ecology and Evolution 9:550.

110. Gottesman S. 2019. Trouble is coming: Signaling pathways that regulate general stress responses in bacteria. The Journal of biological chemistry 294:11685–11700.

111. Davis MC, Kesthely CA, Franklin EA, MacLellan SR. 2017. The essential activities of the bacterial sigma factor. Canadian Journal of Microbiology 63:89–99.

112. Bervoets I, Charlier D. 2019. Diversity, versatility and complexity of bacterial gene regulation mechanisms: opportunities and drawbacks for applications in synthetic biology. FEMS Microbiology Reviews 43:304–339.

113. Tu X, Latifi T, Bougdour A, Gottesman S, Groisman EA. 2006. The PhoP/PhoQ two-component system stabilizes the alternative sigma factor RpoS in Salmonella enterica. Proceedings of the National Academy of Sciences 103:13503–13508.

114. Hengge R. 2008. The Two-Component Network and the General Stress Sigma Factor RpoS (σ<Superscript>S</Superscript>) in <Emphasis Type="Italic">Escherichia coli</Emphasis>. Advances in Experimental Medicine and Biology 631:40–53.

115. Kadowaki T, Yukitake H, Naito M, Sato K, Kikuchi Y, Kondo Y, Shoji M, Nakayama K. 2016. A two-component system regulates gene expression of the type IX secretion component proteins via an ECF sigma factor. Scientific reports 6.

116. Boutte CC, Crosson S. 2013. Bacterial lifestyle shapes the regulation of stringent response activation. Trends in microbiology 21:174.

117. Potrykus K, Cashel M. 2008. (p)ppGpp: still magical? Annual review of microbiology 62:35–51.

118. Das B, Bhadra RK. 2020. (p)ppGpp Metabolism and Antimicrobial Resistance in Bacterial Pathogens. Frontiers in Microbiology 11:2415.

119. Prusa J, Zhu DX, Stallings CL. 2018. The stringent response and Mycobacterium tuberculosis pathogenesis. Pathogens and disease 76.

120. Yang HW, Yu M, Lee JH, Chatnaparat T, Zhao Y. 2020. The stringent response regulator (p) ppGpp mediates virulence gene expression and survival in Erwinia amylovora. BMC Genomics 21:1–19.

121. Schofield WB, Zimmermann-Kogadeeva M, Zimmermann M, Barry NA, Goodman AL. 2018. The Stringent Response Determines the Ability of a Commensal Bacterium to Survive Starvation and to Persist in the Gut. Cell host & microbe 24:120-132.e6.

122. Wu M, McNulty NP, Rodionov DA, Khoroshkin MS, Griffin NW, Cheng J, Latreille P, Kerstetter RA, Terrapon N, Henrissat B, Osterman AL, Gordon JI. 2015. Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut Bacteroides. Science (New York, NY) 350:aac5992.

123. Townsend GE, Han W, Schwalm ND, Hong X, Bencivenga-Barry NA, Goodman AL, Groisman EA, Ii GET, Comstock LE. 2020. A Master Regulator of Bacteroides thetaiotaomicron Gut Colonization Controls Carbohydrate Utilization and an Alternative Protein Synthesis Factor https://doi.org/10.1128/mBio.

124. Mutations Are the Raw Materials of Evolution | Learn Science at Scitable. https://www.nature.com/scitable/knowledge/library/mutations-are-the-raw-materials-of-evolution-17395346/. Retrieved 5 February 2022.

125. Mira A, Ochman H, Moran NA. 2001. Deletional bias and the evolution of bacterial genomes. Trends in genetics : TIG 17:589–596.

126. Morris JJ, Lenski RE, Zinser ER. 2012. The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. mBio 3.

127. Koskiniemi S, Sun S, Berg OG, Andersson DI. 2012. Selection-Driven Gene Loss in Bacteria. PLOS Genetics 8:e1002787.

128. Deszo EL, Steenbergen SM, Freedberg DI, Vimr ER. 2005. Escherichia coli K1 polysialic acid O-acetyltransferase gene, neuO, and the mechanism of capsule form variation involving a mobile contingency locus. Proceedings of the National Academy of Sciences 102:5564–5569.

129. Mordhorst IL, Claus H, Ewers C, Lappann M, Schoen C, Elias J, Batzilla J, Dobrindt U, Wieler LH, Bergfeld AK, Mühlenhoff M, Vogel U. 2009. O-acetyltransferase gene neuO is segregated according to phylogenetic background and contributes to environmental desiccation resistance in Escherichia coli K1. Environmental microbiology 11:3154–3165.

130. Hosseinkhani F, Jabalameli F, Nodeh Farahani N, Taherikalani M, van Leeuwen WB, Emaneini M. 2016. Variable number of tandem repeat profiles and antimicrobial resistance patterns of Staphylococcus haemolyticus strains isolated from blood cultures in children. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases 38:19–21.

131. Paulander W, Andersson DI, Maisnier-Patin S. 2010. Amplification of the gene for isoleucyl-tRNA synthetase facilitates adaptation to the fitness cost of mupirocin resistance in Salmonella enterica. Genetics 185:305–312.

132. Ahmed M, Liang P. 2012. Transposable Elements Are a Significant Contributor to Tandem Repeats in the Human Genome. Comparative and Functional Genomics 2012.

133. Mitkina LN. 2003. Transposition as a way of existence: Phage Mu. Russian Journal of Genetics 39:519–536.

134. Palazzo A, Caizzi R, Viggiano L, Marsano RM. 2017. Does the Promoter Constitute a Barrier in the Horizontal Transposon Transfer Process? Insight from Bari Transposons. Genome Biology and Evolution 9:1637.

135. Kozitskaya S, Cho SH, Dietrich K, Marre R, Naber K, Ziebuhr W. 2004. The Bacterial Insertion Sequence Element IS256 Occurs Preferentially in Nosocomial Staphylococcus epidermidis Isolates: Association with Biofilm Formation and Resistance to Aminoglycosides. Infection and Immunity 72:1210–1215.

136. Soto CY, Menéndez MC, Pérez E, Samper S, Gómez AB, García MJ, Martín C. 2004. IS6110 Mediates Increased Transcription of the phoP Virulence Gene in a Multidrug-Resistant Clinical Isolate Responsible for Tuberculosis Outbreaks. Journal of Clinical Microbiology 42:212.

137. Alonso H, Samper S, Martín C, Otal I. 2013. Mapping IS6110 in high-copy number Mycobacterium tuberculosis strains shows specific insertion points in the Beijing genotype. BMC Genomics 14:1–11.

138. Segrè D, DeLuna A, Church GM, Kishony R. 2005. Modular epistasis in yeast metabolism. Nature genetics 37:77–83.

139. Albert R. 2005. Scale-free networks in cell biology. Journal of cell science 118:4947–4957.

140. Bonhoeffer S, Chappey C, Parkin NT, Whitcomb JM, Petropoulos CJ. 2004. Evidence for positive epistasis in HIV-1. Science (New York, NY) 306:1547–1550.

141. Barabási AL, Oltvai ZN. 2004. Network biology: understanding the cell's functional organization. Nature reviews Genetics 5:101–113.

142. Vogelstein B, Lane D, Levine AJ. 2000. Surfing the p53 network. Nature 408:307–310.

143. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. 1999. From molecular to modular cell biology. Nature 402.

144. Mani R, St Onge RP, Hartman JL th, Giaever G, Roth FP. 2008. Defining genetic interaction. Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 250 Longwood Avenue, Boston, MA 02115, USA. https://doi.org/10.1073/pnas.0712255105.

145. Motter AE, Gulbahce N, Almaas E, Barabási AL. 2008. Predicting synthetic rescues in metabolic networks. Molecular Systems Biology 4:168.

146. Herencias C, Rodríguez-Beltrán J, León-Sampedro R, del Valle AA, Palkovičová J, Cantón R, Millán ÁS. 2021. Collateral sensitivity associated with antibiotic resistance plasmids. eLife 10:1–13.

147. Barbosa C, Roemhild R, Rosenstiel P, Schulenburg H. 2019. Evolutionary stability of collateral sensitivity to antibiotics in the model pathogen Pseudomonas aeruginosa. eLife 8.

148. Motter AE. 2010. Improved network performance via antagonism: From synthetic rescues to multi-drug combinations. BioEssays : news and reviews in molecular, cellular and developmental biology 32:236–245.

149. Baym M, Stone LK, Kishony R. 2016. Multidrug evolutionary strategies to reverse antibiotic resistance. Science (New York, NY) 351.

150. Yin N, Ma W, Pei J, Ouyang Q, Tang C, Lai L. 2014. Synergistic and antagonistic drug combinations depend on network topology. PloS one 9.

151. Zimmer A, Katzir I, Dekel E, Mayo AE, Alon U. 2016. Prediction of multidimensional drug dose responses based on measurements of drug pairs. Proceedings of the National Academy of Sciences of the United States of America 113:10442–10447.

152. Zeng H, Lazarova DL, Bordonaro M. 2014. Mechanisms linking dietary fiber, gut microbiota and colon cancer prevention. World Journal of Gastrointestinal Oncology 6:41.

153. Lee SM, Donaldson GP, Mikulski Z, Boyajian S, Ley K, Mazmanian SK. 2013. Bacterial colonization factors control specificity and stability of the gut microbiota. Nature 2013 501:7467 501:426–429.

154. Donaldson GP, Chou WC, Manson AL, Rogov P, Abeel T, Bochicchio J, Ciulla D, Melnikov A, Ernst PB, Chu H, Giannoukos G, Earl AM, Mazmanian SK. 2020. Spatially distinct physiology of Bacteroides fragilis within the proximal colon of gnotobiotic mice. Nature microbiology 5:746–756.

155. Candela M, Biagi E, Maccaferri S, Turroni S, Brigidi P. 2012. Intestinal microbiota is a plastic factor responding to environmental changes. Trends in Microbiology https://doi.org/10.1016/j.tim.2012.05.003.

156. Sun J, Liao XP, D'Souza AW, Boolchandani M, Li SH, Cheng K, Luis Martínez J, Li L, Feng YJ, Fang LX, Huang T, Xia J, Yu Y, Zhou YF, Sun YX, Deng XB, Zeng ZL, Jiang HX, Fang BH, Tang YZ, Lian XL, Zhang RM, Fang ZW, Yan QL, Dantas G, Liu YH. 2020. Environmental remodeling of human gut microbiota and antibiotic resistome in livestock farms. Nature Communications 2020 11:1 11:1–11.

157. Tsigalou C, Konstantinidis T, Stavropoulou E, Bezirtzoglou EE, Tsakris A. 2020. Potential Elimination of Human Gut Resistome by Exploiting the Benefits of Functional Foods. Frontiers in Microbiology 11:50.

158. Lee K, Kim DW, Lee DH, Kim YS, Bu JH, Cha JH, Thawng CN, Hwang EM, Seong HJ, Sul WJ, Wellington EMH, Quince C, Cha CJ. 2020. Mobile resistome of human gut and pathogen drives anthropogenic bloom of antibiotic resistance. Microbiome 8:1–14.

159. Singh S, Verma N, Taneja N. 2019. The human gut resistome: Current concepts & future prospects. The Indian Journal of Medical Research 150:345.

160. Jain N. 2020. The early life education of the immune system: Moms, microbes and (missed) opportunities. https://doi.org/101080/1949097620201824564 12.

161. Mazmanian SK, Cui HL, Tzianabos AO, Kasper DL. 2005. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. Cell 122:107–118.

162. Mazmanian SK, Round JL, Kasper DL. 2008. A microbial symbiosis factor prevents intestinal inflammatory disease. Nature 453:620–625.

163. Shen Y, Torchia MLG, Lawson GW, Karp CL, Ashwell JD, Mazmanian SK. 2012. Outer membrane vesicles of a human commensal mediate immune regulation and disease protection. Cell Host and Microbe 12:509–520.

164. Surana NK, Kasper DL. 2012. The yin yang of bacterial polysaccharides: Lessons learned from B. fragilis PSA. Immunological Reviews 245:13–26.

165. Blandford LE, Johnston EL, Sanderson JD, Wade WG, Lax AJ. 2019. Promoter orientation of the immunomodulatory Bacteroides fragilis capsular polysaccharide A (PSA) is off in individuals with inflammatory bowel disease (IBD). Gut Microbes 10:569–577.

166. Henke MT, Brown EM, Cassilly CD, Vlamakis H, Xavier RJ, Clardy J. 2021. Capsular polysaccharide correlates with immune response to the human gut microbe Ruminococcus gnavus. Proceedings of the National Academy of Sciences of the United States of America 118.

167. Yue B, Luo X, Yu Z, Mani S, Wang Z, Dou W. 2019. Inflammatory Bowel Disease: A Potential Result from the Collusion between Gut Microbiota and Mucosal Immune System. Microorganisms 7.

168. Aldars-garcía L, Marin AC, Chaparro M, Gisbert JP. 2021. The Interplay between Immune System and Microbiota in Inflammatory Bowel Disease: A Narrative Review. International journal of molecular sciences 22:1–15.

169. Iweala OI, Nagler CR. 2019. The Microbiome and Food Allergy. Annual review of immunology 37:377–403.

170. Eisenstein M. 2020. Microbial ambassadors against food allergies. Nature 588:S11–S13.

171. Vineis JH, Ringus DL, Morrison HG, Delmont TO, Dalal S, Raffals LH, Antonopoulos DA, Rubin DT, Eren AM, Chang EB, Sogin ML. 2016. Patient-specific Bacteroides genome variants in pouchitis. mBio 7.

172. Wu L, Luo Y. 2021. Bacterial Quorum-Sensing Systems and Their Role in Intestinal Bacteria-Host Crosstalk. Frontiers in Microbiology 12:611413.

173. Deng Z, Luo XM, Liu J, Wang H. 2020. Quorum Sensing, Biofilm, and Intestinal Mucosal Barrier: Involvement the Role of Probiotic. Frontiers in Cellular and Infection Microbiology 10:504.

174. Rakoff-Nahoum S, Foster KR, Comstock LE. 2016. The evolution of cooperation within the gut microbiota. Nature 533:255.

175. Coyte KZ, Rakoff-Nahoum S. 2019. Understanding Competition and Cooperation within the Mammalian Gut Microbiome. Current biology : CB 29:R538.

176. Sachs JL, Mueller UG, Wilcox TP, Bull JJ. 2004. The evolution of cooperation. The Quarterly review of biology 79:135–160.

177. Foster KR, Wenseleers T. 2006. A general model for the evolution of mutualisms. Journal of evolutionary biology 19:1283–1293.

178. Frank SA. 2010. A general model of the public goods dilemma. Journal of evolutionary biology 23:1245–1250.

179. Mitri S, Richard Foster K. 2013. The genotypic view of social interactions in microbial communities. Annual review of genetics 47:247–273.

180. Chatzidaki-Livanis M, Geva-Zatorsky N, Comstock LE, Hooper L v. 2016. Bacteroides fragilis type VI secretion systems use novel effector and immunity proteins to antagonize human gut Bacteroidales species. Proceedings of the National Academy of Sciences of the United States of America 113:3627–3632.

181. Verster AJ, Ross BD, Radey MC, Bao Y, Goodman AL, Mougous JD, Borenstein E. 2017. The Landscape of Type VI Secretion across Human Gut Microbiomes Reveals Its Role in Community Composition. Cell Host and Microbe 22:411-419.e4.

182. Coyne MJ, Comstock LE. 2019. Type VI Secretion Systems and the Gut Microbiota. Microbiology Spectrum 7.

183. García-Bayona L, Coyne MJ, Comstock LE. 2021. Mobile Type VI secretion system loci of the gut Bacteroidales display extensive intra-ecosystem transfer, multi-species spread and geographical clustering. PLoS Genetics 17.

184. Scheuerl T, Hopkins M, Nowell RW, Rivett DW, Barraclough TG, Bell T. 2020. Bacterial adaptation is constrained in complex communities. Nature Communications 11.

185. Morgan MC, Boyette M, Goforth C, Sperry KV, Greene SR. 2009. Comparison of the Biolog OmniLog Identification System and 16S ribosomal RNA gene sequencing for accuracy in identification of atypical bacteria of clinical origin. Journal of microbiological methods 79:336–343.

186. Liu H, Price MN, Waters RJ, Ray J, Carlson HK, Lamson JS, Chakraborty R, Arkin AP, Deutschbauer AM. 2018. Magic Pools: Parallel Assessment of Transposon Delivery Vectors in Bacteria. mSystems 3.

187. Gallagher LA, Ramage E, Jacobs MA, Kaul R, Brittnacher M, Manoil C. 2007. A comprehensive transposon mutant library of Francisella novicida, a bioweapon surrogate. National Acad Sciences.

188. Liberati NT, Urbach JM, Miyata S, Lee DG, Drenkard E, Wu G, Villanueva J, Wei T, Ausubel FM. 2006. An ordered, nonredundant library of Pseudomonas aeruginosa strain PA14 transposon insertion mutants. National Acad Sciences.

189. Lewenza S, Falsafi RK, Winsor G, Gooderham WJ, Mcphee JB, Brinkman FSL, Hancock REW. Construction of a mini-Tn5-luxCDABE mutant library in Pseudomonas aeruginosa PAO1: a tool for identifying differentially regulated genes. genome.cshlp.org https://doi.org/10.1101/gr.3513905.

190. Jacobs MA, Alwood A, Thaipisuttikul I, Spencer D, Haugen E, Ernst S, Will O, Kaul R, Raymond C, Levy R, Chun-Rong L, Guenthner D, Bovee D, Olson M v, Manoil C. 2003. Comprehensive transposon mutant library of Pseudomonas aeruginosa. National Acad Sciences.

191. Lam LH, Monack DM. 2014. Intraspecies Competition for Niches in the Distal Gut Dictate Transmission during Persistent Salmonella Infection. PLOS Pathogens 10:e1004527.

192. Rooks MG, Garrett WS. 2016. Gut microbiota, metabolites and host immunity. Nature reviews Immunology 16:341.

193. Alseekh S, Aharoni A, Brotman Y, Contrepois K, D'Auria J, Ewald J, C. Ewald J, Fraser PD, Giavalisco P, Hall RD, Heinemann M, Link H, Luo J, Neumann S, Nielsen J, Perez de Souza L, Saito K, Sauer U, Schroeder FC, Schuster S, Siuzdak G, Skirycz A, Sumner LW, Snyder MP, Tang H, Tohge T, Wang Y, Wen W, Wu S, Xu G, Zamboni N, Fernie AR. 2021. Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices. Nature Methods 2021 18:7 18:747–756.

194. Misra BB, van der Hooft JJJ. 2016. Updates in metabolomics tools and resources: 2014–2015. ELECTROPHORESIS 37:86–110.

195. Wong AC, Levy M. 2019. New Approaches to Microbiome-Based Therapies. mSystems 4.

196. Shi X, Zhang J, Mo L, Shi J, Qin M, Huang X. 2019. Efficacy and safety of probiotics in eradicating Helicobacter pylori: A network meta-analysis. Medicine 98:e15180.

197. Hosseini M, Yousefifard M, Ataei N, Oraii A, Mirzay Razaz J, Izadi A. 2017. The efficacy of probiotics in prevention of urinary tract infection in children: A systematic review and meta-analysis. Journal of Pediatric Urology 13:581–591.

198. Seminario-Amez M, López-López J, Estrugo-Devesa A, Ayuso-Montero R, Jané-Salas E. 2017. Probiotics and oral health: A systematic review. Medicina Oral, Patologia Oral y Cirugia Bucal 22:e282–e288.

199. Dixon A, Robertson K, Yung A, Que M, Randall H, Wellalagodage D, Cox T, Robertson D, Chi C, Sun J. 2020. Efficacy of Probiotics in Patients of Cardiovascular Disease Risk: A Systematic Review and Meta-analysis. Current Hypertension Reports 22.

200. Redman MG, Ward EJ, Phillips RS. 2014. The efficacy and safety of probiotics in people with cancer: A systematic review. Annals of Oncology 25:1919–1929.

201. Dang X, Xu M, Liu D, Zhou D, Yang W. 2020. Assessing the efficacy and safety of fecal microbiota transplantation and probiotic VSL#3 for active ulcerative colitis: A systematic review and meta-analysis. PLoS ONE 15.

202. Xu D, Chen VL, Steiner CA, Berinstein JA, Eswaran S, Waljee AK, Higgins PDR, Owyang C. 2019. Efficacy of Fecal Microbiota Transplantation in Irritable Bowel Syndrome: A Systematic Review and Meta-Analysis. American Journal of Gastroenterology 114:1043–1050.

203. Markowiak P, Ślizewska K. 2017. Effects of probiotics, prebiotics, and synbiotics on human health. Nutrients 9.

204. Green JE, Davis JA, Berk M, Hair C, Loughman A, Castle D, Athan E, Nierenberg AA, Cryan JF, Jacka F, Marx W. 2020. Efficacy and safety of fecal microbiota transplantation for the treatment of diseases other than Clostridium difficile infection: a systematic review and meta-analysis. Gut Microbes 12:1–25.

205. Wang JW, Kuo CH, Kuo FC, Wang YK, Hsu WH, Yu FJ, Hu HM, Hsu PI, Wang JY, Wu DC. 2019. Fecal microbiota transplantation: Review and update. Journal of the Formosan Medical Association 118:S23–S31.

206. Suez J, Zmora N, Zilberman-Schapira G, Mor U, Dori-Bachash M, Bashiardes S, Zur M, Regev-Lehavi D, Ben-Zeev Brik R, Federici S, Horn M, Cohen Y, Moor AE, Zeevi D, Korem T, Kotler E, Harmelin A, Itzkovitz S, Maharshak N, Shibolet O, Pevsner-Fischer M, Shapiro H, Sharon I, Halpern Z, Segal E, Elinav E. 2018. Post-Antibiotic Gut Mucosal Microbiome Reconstitution Is Impaired by Probiotics and Improved by Autologous FMT. Cell 174:1406-1423.e16.

207. Szathmáry E, Jordán F, Pál C. 2001. Molecular biology and evolution. Can genes explain biological complexity? Collegium Budapest (Institute for Advanced Study), 2 Szentháromság u., H-1014 Budapest, Hungary. szathmary\char64colbud.hu https://doi.org/10.1126/science.1060852.

208. Tong AHY, Lesage G, Bader GD, Ding H, Xu H, Xin X, Young J, Berriz GF, Brost RL, Chang M, Chen YQ, Cheng X, Chua G, Friesen H, Goldberg DS, Haynes J, Humphries C, He G, Hussein S, Ke L, Krogan N, Li Z, Levinson JN, Lu H, Ménard P, Munyana C, Parsons AB, Ryan O, Tonikian R, Roberts T, Sdicu AM, Shapiro J, Sheikh B, Suter B, Wong SL, Zhang L v., Zhu H, Burd CG, Munro S, Sander C, Rine J, Greenblatt J, Peter M, Bretscher A, Bell G, Roth FP, Brown GW, Andrews B, Bussey H, Boone C. 2004. Global mapping of the yeast genetic interaction network. Science (New York, NY) 303:808–813.

209. Butland G, Babu M, Díaz-Mejía JJ, Bohdana F, Phanse S, Gold B, Yang W, Li J, Gagarinova AG, Pogoutse O, Mori H, Wanner BL, Lo H, Wasniewski J, Christopolous C, Ali M, Venn P, Safavi-Naini A, Sourour N, Caron S, Choi JY, Laigle L, Nazarians-Armavil A, Deshpande A, Joe S, Datsenko KA, Yamamoto N, Andrews BJ, Boone C, Ding H, Sheikh B, Moreno-Hagelseib G, Greenblatt JF, Emili A. 2008. eSGA: E. coli synthetic genetic array analysis. Nat Methods2008/08/05. 5:789–95.

210. Typas A, Nichols RJ, Siegele DA, Shales M, Collins SR, Lim B, Braberg H, Yamamoto N, Takeuchi R, Wanner BL, Mori H, Weissman JS, Krogan NJ, Gross CA. 2008. High-throughput, quantitative analyses of genetic interactions in E. coli. Department of Microbiology and Immunology, University of California at San Francisco, 600 16th Street, San Francisco, California 94158, USA. https://doi.org/10.1038/nmeth.1240.

211. Costanzo M, VanderSluis B, Koch EN, Baryshnikova A, Pons C, Tan G, Wang W, Usaj M, Hanchard J, Lee SD, Pelechano V, Styles EB, Billmann M, van Leeuwen J, van Dyk N, Lin ZY, Kuzmin E, Nelson J, Piotrowski JS, Srikumar T, Bahr S, Chen Y, Deshpande R, Kurat CF, Li SC, Li Z, Usaj MM, Okada H, Pascoe N, Luis BJS, Sharifpoor S, Shuteriqi E, Simpkins SW, Snider J, Suresh HG, Tan Y, Zhu H, Malod-Dognin N, Janjic V, Przulj N, Troyanskaya OG, Stagljar I, Xia T, Ohya Y, Gingras AC, Raught B, Boutros M, Steinmetz LM, Moore CL, Rosebrock AP, Caudy AA, Myers CL, Andrews B, Boone C. 2016. A global genetic interaction network maps a wiring diagram of cellular function. Science (New York, NY) 353.

212. van Opijnen T, Bodi KL, Camilli A. 2009. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nature methods 6:767–772.

213. Ghim CM, Goh KI, Kahng B. 2005. Lethality and synthetic lethality in the genome-wide metabolic network of Escherichia coli. School of Physics, Seoul National University NS50, Seoul 151-747, Republic of Korea. https://doi.org/10.1016/j.jtbi.2005.04.025.

214. Shlomi T, Berkman O, Ruppin E. 2005. Regulatory on/off minimization of metabolic flux changes after genetic perturbations. School of Computer Science, , Tel Aviv University, Tel Aviv 69978, Israel. https://doi.org/10.1073/pnas.0406346102.

215. Aziz RK, Monk JM, Lewis RM, In Loh S, Mishra A, Abhay Nagle A, Satyanarayana C, Dhakshinamoorthy S, Luche M, Kitchen DB, Andrews KA, Fong NL, Li HJ, Palsson BO, Charusanti P. 2015. Systems biology-guided identification of synthetic lethal gene pairs and its potential use to discover antibiotic combinations. Scientific Reports 5:16025.

216. Wood KB, Cluzel P. 2012. Trade-offs between drug toxicity and benefit in the multi-antibiotic resistance system underlie optimal growth of E. coli. BMC systems biology 6.

217. Imamovic L, Sommer MO. 2013. Use of collateral sensitivity networks to design drug cycling protocols that avoid resistance development. Department of Systems Biology, Technical University of Denmark, DK-2800 Lyngby, Denmark. https://doi.org/10.1126/scitranslmed.3006609.

218. Lehner B. 2011. Molecular mechanisms of epistasis within and between genes. Trends in genetics : TIG 27:323–331.

219. Snitkin ES, Segrè D. 2011. Epistatic interaction maps relative to multiple metabolic phenotypes. PLoS genetics 7.

220. Reguly T, Breitkreutz A, Boucher L, Breitkreutz BJ, Hon GC, Myers CL, Parsons A, Friesen H, Oughtred R, Tong A, Stark C, Ho Y, Botstein D, Andrews B, Boone C,

Troyanskya OG, Ideker T, Dolinski K, Batada NN, Tyers M. 2006. Comprehensive curation and analysis of global interaction networks in Saccharomyces cerevisiae. Journal of biology 5.

221. Motter AE, Gulbahce N, Almaas E, Barabási AL. 2008. Predicting synthetic rescues in metabolic networks. Molecular systems biology 4.

222. McCloskey D, Xu S, Sandberg TE, Brunk E, Hefner Y, Szubin R, Feist AM, Palsson BO. 2018. Evolution of gene knockout strains of E. coli reveal regulatory architectures governed by metabolism. Nature communications 9.

223. Wytock TP, Fiebig A, Willett JW, Herrou J, Fergin A, Motter AE, Crosson S. 2018. Experimental evolution of diverse Escherichia coli metabolic mutants identifies genetic loci for convergent adaptation of growth rate. PLoS genetics 14.

224. Weinreich DM, Watson RA, Chao L. 2005. Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA. dmw\char64post.harvard.edu.

225. Fraenkel DG. 1987. Glycolysis, pentose phosphate pathway, and Entner-Doudoroff pathway. Washington, D.C.

226. Ashworth A, Lord CJ, Reis-Filho JS. 2011. Genetic interactions in cancer progression and treatment. The Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, Fulham Road, London SW3 6JB, UK. alan.ashworth\char64icr.ac.uk https://doi.org/10.1016/j.cell.2011.03.020.

227. Gerdes K, Maisonneuve E. 2012. Bacterial persistence and toxin-antitoxin loci. Annual review of microbiology 66:103–123.

228. Farha MA, Czarny TL, Myers CL, Worrall LJ, French S, Conrady DG, Wang Y, Oldfield E, Strynadka NCJ, Brown ED. 2015. Antagonism screen for inhibitors of bacterial cell wall biogenesis uncovers an inhibitor of undecaprenyl diphosphate synthase. Proceedings of the National Academy of Sciences of the United States of America 112:11048–11053.

229. Brown ED, Wright GD. 2016. Antibacterial drug discovery in the resistance era. Nature2016/01/23. 529:336–43.

230. Tamer YT, Gaszek IK, Abdizadeh H, Batur TA, Reynolds KA, Atilgan AR, Atilgan C, Toprak E. 2019. High-Order Epistasis in Catalytic Power of Dihydrofolate Reductase Gives Rise to a Rugged Fitness Landscape in the Presence of Trimethoprim Selection. Molecular biology and evolution 36:1533–1550.

231. Cornelius SP, Lee JS, Motter AE. 2011. Dispensability of Escherichia coli's latent pathways. Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA. https://doi.org/10.1073/pnas.1009772108.

232. González C, Ray JCJ, Manhart M, Adams RM, Nevozhay D, Morozov A v, Balázsi G. 2015. Stress-response balance drives the evolution of a network module and its host genome. Molecular systems biology 11:827.

233. Kheir Gouda M, Manhart M, Balázsi G. 2019. Evolutionary regain of lost gene circuit function. Proc Natl Acad Sci U S A2019/11/23. 116:25162–25171.

234. Pena-Miller R, Laehnemann D, Jansen G, Fuentes-Hernandez A, Rosenstiel P, Schulenburg H, Beardmore R. 2013. When the most potent combination of antibiotics selects for the greatest bacterial load: the smile-frown transition. PLoS biology 11.

235. Iram S, Dolson E, Chiel J, Pelesko J, Krishnan N, Güngör Ö, Kuznets-Speck B, Deffner S, Ilker E, Scott JG, Hinczewski M. 2020. Controlling the speed and trajectory of evolution with counterdiabatic driving. Nature Physics 2020 17:1 17:135–142.

236. Maltas J, Wood KB. 2019. Pervasive and diverse collateral sensitivity profiles inform optimal strategies to limit antibiotic resistance. PLoS biology 17.

237. Nichol D, Rutter J, Bryant C, Hujer AM, Lek S, Adams MD, Jeavons P, Anderson ARA, Bonomo RA, Scott JG. 2019. Antibiotic collateral sensitivity is contingent on the repeatability of evolution. Nature communications 10.

238. Dean Z, Maltas J, Wood KB. 2020. Antibiotic interactions shape short-term evolution of resistance in E. faecalis. PLoS pathogens 16.

239. Lindsey HA, Gallie J, Taylor S, Kerr B. 2013. Evolutionary rescue from extinction is contingent on a lower rate of environmental change. Nature 494:463–467.

240. Enne VI, Delsol AA, Roe JM, Bennett PM. 2004. Rifampicin resistance and its fitness cost in Enterococcus faecium. The Journal of antimicrobial chemotherapy 53:203–207.

241. Paulander W, Maisnier-Patin S, Andersson DI. 2009. The fitness cost of streptomycin resistance depends on rpsL mutation, carbon source and RpoS (sigmaS). Genetics 183:539–546.

242. Maharjan R, Ferenci T. 2017. The fitness costs and benefits of antibiotic resistance in drug-free microenvironments encountered in the human body. Environmental microbiology reports 9:635–641.

243. Bollenbach T, Quan S, Chait R, Kishony R. 2009. Nonoptimal microbial response to antibiotics underlies suppressive drug interactions. Cell 139:707–718.

244. Quan S, Ray JCJ, Kwota Z, Duong T, Balázsi G, Cooper TF, Monds RD. 2012. Adaptive evolution of the lactose utilization network in experimentally evolved populations of Escherichia coli. PLoS genetics 8.

245. Roemhild R, Gokhale CS, Dirksen P, Blake C, Rosenstiel P, Traulsen A, Andersson DI, Schulenburg H. 2018. Cellular hysteresis as a principle to maximize the efficacy of

antibiotic therapy. Proceedings of the National Academy of Sciences of the United States of America 115:9767–9772.

246. Blagosklonny M v. 1999. Drug-resistance enables selective killing of resistant leukemia cells: exploiting of drug resistance instead of reversal. Medicine Branch, National Cancer Institute, Building 10, R 12N226, NIH, Bethesda, MD 20892, USA. https://doi.org/10.1038/sj.leu.2401623.

247. Chait R, Craney A, Kishony R. 2007. Antibiotic interactions that select against resistance. Nature 446:668–671.

248. Palmer AC, Chait R, Kishony R. 2018. Nonoptimal Gene Expression Creates Latent Potential for Antibiotic Resistance. Molecular biology and evolution 35:2669–2684.

249. Lázár V, Nagy I, Spohn R, Csörgo ' B, Györkei Á, Nyerges Á, Horváth B, Vörös A, Busa-Fekete R, Hrtyan M, Bogos B, Méhi O, Fekete G, Szappanos B, Kégl B, Papp B, Pál C. 2014. Genome-wide analysis captures the determinants of the antibiotic cross-resistance interaction network. Nature communications 5.

250. Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M, Galagan J, Niemann S, Gagneux S. 2011. Whole-genome sequencing of rifampicin-resistant Mycobacterium tuberculosis strains identifies compensatory mutations in RNA polymerase genes. Nature genetics 44:106–110.

251. Stefan MA, Ugur FS, Garcia GA. 2018. Source of the Fitness Defect in Rifamycin-Resistant Mycobacterium tuberculosis RNA Polymerase and the Mechanism of Compensation by Mutations in the β' Subunit. Antimicrobial agents and chemotherapy 62.

252. Wang B, Zhang X, Yu X, Cui Z, Wang Z, Chen T, Zhao X. 2019. Evolutionary engineering of Escherichia coli for improved anaerobic growth in minimal medium accelerated lactate production. Applied microbiology and biotechnology 103:2155–2170.

253. Campbell EA, Korzheva N, Mustaev A, Murakami K, Nair S, Goldfarb A, Darst SA. 2001. Structural mechanism for rifampicin inhibition of bacterial rna polymerase. Cell 104:901–912.

254. Zhou YN, Lubkowska L, Hui M, Court C, Chen S, Court DL, Strathern J, Jin DJ, Kashlev M. 2013. Isolation and characterization of RNA polymerase rpoB mutations that alter transcription slippage during elongation in Escherichia coli. The Journal of biological chemistry 288:2700–2710.

255. Trautinger BW, Lloyd RG. 2002. Modulation of DNA repair by mutations flanking the DNA channel through RNA polymerase. The EMBO journal 21:6944–6953.

256. Jin DJ, Walter WA, Gross CA. 1988. Characterization of the termination phenotypes of rifampicin-resistant mutants. Journal of molecular biology 202:245–253.

257. Garibyan L, Huang T, Kim M, Wolff E, Nguyen A, Nguyen T, Diep A, Hu K, Iverson A, Yang H, Miller JH. 2003. Use of the rpoB gene to determine the specificity of base substitution mutations on the Escherichia coli chromosome. DNA repair 2:593–608.

258. Jin DJ, Gross CA. 1988. Mapping and sequencing of mutations in the Escherichia coli rpoB gene that lead to rifampicin resistance. Journal of molecular biology 202:45–58.

259. Ederth J, Mooney RA, Isaksson LA, Landick R. 2006. Functional interplay between the jaw domain of bacterial RNA polymerase and allele-specific residues in the product RNA-binding pocket. Journal of molecular biology 356:1163–1179.

260. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. 2019. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Division of Bioinformatics, Department of Preventive Medicine, Keck School of Medicine of USC, University of Southern California, Los Angeles, CA 90033, USA. School of Life Sciences, Guangzhou University, Guangzhou 510006, China. https://doi.org/10.1093/nar/gky1038.

261. Galperin MY, Makarova KS, Wolf YI, Koonin E v. 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic acids research 43:D261–D269.

262. Sastry A v., Gao Y, Szubin R, Hefner Y, Xu S, Kim D, Choudhary KS, Yang L, King ZA, Palsson BO. 2019. The Escherichia coli transcriptome mostly consists of independently regulated modules. Nature communications 10.

263. Balázsi G, Barabási AL, Oltvai ZN. 2005. Topological units of environmental signal processing in the transcriptional regulatory network of Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America 102:7841–7846.

264. Santos-Zavaleta A, Salgado H, Gama-Castro S, Sánchez-Pérez M, Gómez-Romero L, Ledezma-Tejeida D, García-Sotelo JS, Alquicira-Hernández K, Muñiz-Rascado LJ, Peña-Loredo P, Ishida-Gutiérrez C, Velázquez-Ramírez DA, del Moral-Chávez V, Bonavides-Martínez C, Méndez-Cruz CF, Galagan J, Collado-Vides J. 2019. RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in E. coli K-12. Nucleic acids research 47:D212–D220.

265. Ross W, Vrentas CE, Sanchez-Vazquez P, Gaal T, Gourse RL. 2013. The magic spot: a ppGpp binding site on E. coli RNA polymerase responsible for regulation of transcription initiation. Molecular cell 50:420–429.

266. Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson JR, Hwa T. 2015. Overflow metabolism in Escherichia coli results from efficient proteome allocation. Nature 528:99–104.

267. Hall AR, Iles JC, MacLean RC. 2011. The fitness cost of rifampicin resistance in Pseudomonas aeruginosa depends on demand for RNA polymerase. Genetics 187:817–822.

268. Vvedenskaya IO, Vahedian-Movahed H, Bird JG, Knoblauch JG, Goldman SR, Zhang Y, Ebright RH, Nickels BE. 2014. Interactions between RNA polymerase and the "core recognition element" counteract pausing. Science (New York, NY) 344:1285–1289.

269. di Salvo M, Puccio S, Peano C, Lacour S, Alifano P. 2019. RhoTermPredict: an algorithm for predicting Rho-dependent transcription terminators based on Escherichia coli, Bacillus subtilis and Salmonella enterica databases. BMC bioinformatics 20.

270. Barrick JE, Kauth MR, Strelioff CC, Lenski RE. 2010. Escherichia coli rpoB mutants have increased evolvability in proportion to their fitness defects. Molecular biology and evolution 27:1338–1347.

271. Qi Q, Toll-Riera M, Heilbron K, Preston GM, Maclean RC. 2016. The genomic basis of adaptation to the fitness cost of rifampicin resistance in Pseudomonas aeruginosa. Proceedings Biological sciences 283.

272. Hietpas RT, Bank C, Jensen JD, Bolon DNA. 2013. Shifting fitness landscapes in response to altered environments. Evolution; international journal of organic evolution 67:3512–3522.

273. Blanquart F, Achaz G, Bataillon T, Tenaillon O. 2014. Properties of selected mutations and genotypic landscapes under Fisher's geometric model. Evolution; international journal of organic evolution 68:3537–3554.

274. Clarke L, Pelin A, Phan M, Wong A. 2020. The effect of environmental heterogeneity on the fitness of antibiotic resistance mutations in Escherichia coli https://doi.org/10.1007/s10682-019-10027-y.

275. Trindade S, Sousa A, Xavier KB, Dionisio F, Ferreira MG, Gordo I. 2009. Positive epistasis drives the acquisition of multidrug resistance. Instituto Gulbenkian de Ciência, Oeiras, Portugal. https://doi.org/10.1371/journal.pgen.1000578.

276. Kishony R, Leibler S. 2003. Environmental stresses can alleviate the average deleterious effect of mutations. Journal of biology 2.

277. Conrad TM, Frazier M, Joyce AR, Cho BK, Knight EM, Lewis NE, Landick R, Palsson B. 2010. RNA polymerase mutants found through adaptive evolution reprogram Escherichia coli for optimal growth in minimal media. Proceedings of the National Academy of Sciences of the United States of America 107:20500–20505.

278. Hall AE, Karkare K, Cooper VS, Bank C, Cooper TF, Moore FBG. 2019. Environment changes epistasis to alter trade-offs along alternative evolutionary paths. Evolution; international journal of organic evolution 73:2094–2105.

279. Angst DC, Hall AR. 2013. The cost of antibiotic resistance depends on evolutionary history in Escherichia coli. Institute of Integrative Biology, ETH Zürich, Zürich, CH-8092, Switzerland. https://doi.org/10.1186/1471-2148-13-163.

280. Hall AR. 2013. Genotype-by-environment interactions due to antibiotic resistance and adaptation in Escherichia coli. Journal of evolutionary biology 26:1655–1664.

281. Sabarly V, Aubron C, Glodt J, Balliau T, Langella O, Chevret D, Rigal O, Bourgais A, Picard B, de Vienne D, Denamur E, Bouvet O, Dillmann C. 2016. Interactions between genotype and environment drive the metabolic phenotype within Escherichia coli isolates. Environmental microbiology 18:100–117.

282. Maltas J, Krasnick B, Wood KB. 2020. Using Selection by Nonantibiotic Stressors to Sensitize Bacteria to Antibiotics. Department of Biophysics, University of Michigan, Ann Arbor, MI. Department of Physics, University of Michigan, Ann Arbor, MI. https://doi.org/10.1093/molbev/msz303.

283. Landick R. 2006. The regulatory roles and mechanism of transcriptional pausing. Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53705, USA. rlandick\char64wisc.edu https://doi.org/10.1042/bst0341062.

284. Klumpp S, Hwa T. 2008. Stochasticity and traffic jams in the transcription of ribosomal RNA: Intriguing role of termination and antitermination. Center for Theoretical Biological Physics and Department of Physics, University of California at San Diego, La Jolla, CA 92093-0374, USA. klumpp\char64ctbp.ucsd.edu https://doi.org/10.1073/pnas.0806084105.

285. Roux D, Danilchanka O, Guillard T, Cattoir V, Aschard H, Fu Y, Angoulvant F, Messika J, Ricard JD, Mekalanos JJ, Lory S, Pier GB, Skurnik D. 2015. Fitness cost of antibiotic susceptibility during bacterial infection. Science translational medicine 7.

286. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA. https://doi.org/10.1073/pnas.120163297.

287. Lee H, Popodi E, Tang H, Foster PL. 2012. Rate and molecular spectrum of spontaneous mutations in the bacterium Escherichia coli as determined by whole-genome sequencing. Proceedings of the National Academy of Sciences of the United States of America 109.

288. Deatherage DE, Barrick JE. 2014. Identification of Mutations in Laboratory-Evolved Microbes from Next-Generation Sequencing Data Using breseq, p. 165–188. *In* Sun, L, Shou, W (eds.), Engineering and Analyzing Multicellular Systems: Methods and Protocols. Springer New York, New York, NY.

289. Zhu A, Ibrahim JG, Love MI. 2019. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Department of Biostatistics, University of North Carolina-Chapel Hill, NC, USA. Department of Genetics, University of North Carolina-Chapel Hill, NC, USA. https://doi.org/10.1093/bioinformatics/bty895.

290. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. 2001. Controlling the false discovery rate in behavior genetics research. Department of Statistics and O.R., The Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. ybenja\char64post.tau.ac.il https://doi.org/10.1016/s0166-4328(01)00297-2.

291. Durfee T, Hansen AM, Zhi H, Blattner FR, Ding JJ. 2008. Transcription profiling of the stringent response in Escherichia coli. Journal of bacteriology 190:1084–1096.

292. O'Malley MA. 2007. The nineteenth century roots of "everything is everywhere." Nature Reviews Microbiology 2007 5:8 5:647–651.

293. Erickson KE, Otoupal PB, Chatterjee A. 2015. Gene Expression Variability Underlies Adaptive Resistance in Phenotypically Heterogeneous Bacterial Populations. ACS Infectious Diseases 1:555–567.

294. Veening J-W, Smits WK, Kuipers OP. 2008. Bistability, Epigenetics, and Bet-Hedging in Bacteria. http://dx.doi.org/101146/annurev.micro62081307163002 62:193–210.

295. Ackermann M. 2015. A functional perspective on phenotypic heterogeneity in microorganisms. Nature Reviews Microbiology 2015 13:8 13:497–508.

296. Cao M, Goodrich-Blair H. 2017. Ready or Not: Microbial Adaptive Responses in Dynamic Symbiosis Environments. Journal of Bacteriology 199.

297. Goodman AL, McNulty NP, Zhao Y, Leip D, Mitra RD, Lozupone CA, Knight R, Gordon JI. 2009. Identifying Genetic Determinants Needed to Establish a Human Gut Symbiont in Its Habitat. Cell Host and Microbe 6:279–289.

298. Wu M, McNulty NP, Rodionov DA, Khoroshkin MS, Griffin NW, Cheng J, Latreille P, Kerstetter RA, Terrapon N, Henrissat B, Osterman AL, Gordon JI. 2015. Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut Bacteroides. Science 350.

299. Liu H, Shiver AL, Price MN, Carlson HK, Trotter V V., Chen Y, Escalante V, Ray J, Hern KE, Petzold CJ, Turnbaugh PJ, Huang KC, Arkin AP, Deutschbauer AM. 2021. Functional genetics of human gut commensal Bacteroides thetaiotaomicron reveals metabolic requirements for growth across environments. Cell Reports 34.

300. Suez J, Zmora N, Zilberman-Schapira G, Mor U, Dori-Bachash M, Bashiardes S, Zur M, Regev-Lehavi D, Ben-Zeev Brik R, Federici S, Horn M, Cohen Y, Moor AE, Zeevi D, Korem T, Kotler E, Harmelin A, Itzkovitz S, Maharshak N, Shibolet O, Pevsner-Fischer M, Shapiro H, Sharon I, Halpern Z, Segal E, Elinav E. 2018. Post-Antibiotic Gut Mucosal Microbiome Reconstitution Is Impaired by Probiotics and Improved by Autologous FMT. Cell 174:1406-1423.e16.

301. Sommer F, Nookaew I, Sommer N, Fogelstrand P, Bäckhed F. 2015. Site-specific programming of the host epithelial transcriptome by the gut microbiota. Genome Biology 2015 16:1 16:1–15.

302. Arora T, Akrami R, Pais R, Bergqvist L, Johansson BR, Schwartz TW, Reimann F, Gribble FM, Bäckhed F. 2018. Microbial regulation of the L cell transcriptome. Scientific Reports 2018 8:1 8:1–9.

303. Wetmore KM, Price MN, Waters RJ, Lamson JS, He J, Hoover CA, Blow MJ, Bristow J, Butland G, Arkin AP, Deutschbauer A. 2015. Rapid Quantification of Mutant Fitness in Diverse Bacteria by Sequencing Randomly Bar-Coded Transposons. mBio 6:e00306-15.

304. Porter NT, Canales P, Peterson DA, Martens EC. 2017. A Subset of Polysaccharide Capsules in the Human Symbiont Bacteroides thetaiotaomicron Promote Increased Competitive Fitness in the Mouse Gut. Cell Host and Microbe 22:494-506.e8.

305. Peterson DA, McNulty NP, Guruge JL, Gordon JI. 2007. IgA Response to Symbiotic Bacteria as a Mediator of Gut Homeostasis. Cell Host and Microbe 2:328–339.

306. Goodman AL, McNulty NP, Zhao Y, Leip D, Mitra RD, Lozupone CA, Knight R, Gordon JI. 2009. Identifying Genetic Determinants Needed to Establish a Human Gut Symbiont in Its Habitat. Cell Host and Microbe 6:279–289.

307. Watson AR, Füssel J, Veseli I, Zaal DeLongchamp J, Silva M, Trigodet F, Lolans K, Shaiber A, Fogarty E, Runde M, Quince C, Yu MK, Söylev A, Morrison HG, Lee ST, Kao D, Rubin DT, Jabri B, Louie T, Murat Eren A, Bay Paul Center J. Adaptive ecological processes and metabolic independence drive microbial colonization and resilience in the human gut https://doi.org/10.1101/2021.03.02.433653.

308. Martens EC, Chiang HC, Gordon JI. 2008. Mucosal Glycan Foraging Enhances Fitness and Transmission of a Saccharolytic Human Gut Bacterial Symbiont. Cell Host and Microbe 4:447–457.

309. Vasquez KS, Willis L, Cira NJ, Ng KM, Pedro MF, Aranda-Díaz A, Rajendram M, Yu FB, Higginbottom SK, Neff N, Sherlock G, Xavier KB, Quake SR, Sonnenburg JL, Good BH, Huang KC. 2021. Quantifying rapid bacterial evolution and transmission within the mouse intestine. Cell Host and Microbe 29:1454-1468.e4.

310. Liu H, Shiver AL, Price MN, Carlson HK, Trotter V v, Chen Y, Escalante V, Ray J, Hern KE, Petzold CJ, Turnbaugh PJ, Huang KC, Arkin AP, Deutschbauer AM. 2021. Functional genetics of human gut commensal Bacteroides thetaiotaomicron reveals metabolic requirements for growth across environments. Cell Reports 34.

311. Chandler M, Fayet O, Rousseau P, Ton Hoang B, Duval-Valentin G. 2015. Copy-out–Paste-in Transposition of IS 911 : A Major Transposition Pathway . Microbiology Spectrum 3.

312. Wu M, McNulty NP, Rodionov DA, Khoroshkin MS, Griffin NW, Cheng J, Latreille P, Kerstetter RA, Terrapon N, Henrissat B, Osterman AL, Gordon JI. 2015. Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut Bacteroides. Science 350.

313. Sonnenburg JL, Xu J, Leip DD, Chen C-H, Westover BP, Weatherford J, Buhler JD, Gordon JI. 2005. Glycan Foraging in Vivo by an Intestine-Adapted Bacterial Symbiont. Science 307:1955–1959.

314. Schofield WB, Zimmermann-Kogadeeva M, Zimmermann M, Barry NA, Goodman AL. 2018. The Stringent Response Determines the Ability of a Commensal Bacterium to Survive Starvation and to Persist in the Gut. Cell Host and Microbe 24:120-132.e6.

315. Magnúsdóttir S, Ravcheev D, de Crécy-Lagard V, Thiele I. 2015. Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. Frontiers in Genetics 6.

316. Carfrae LA, MacNair CR, Brown CM, Tsai CN, Weber BS, Zlitni S, Rao VN, Chun J, Junop MS, Coombes BK, Brown ED. 2020. Mimicking the human environment in mice reveals that inhibiting biotin biosynthesis is effective against antibiotic-resistant pathogens. Nature Microbiology. Nature Research https://doi.org/10.1038/s41564-019-0595-2.

317. Crofts AA, Giovanetti SM, Rubin EJ, Poly FM, Gutiérrez RL, Talaat KR, Porter CK, Riddle MS, DeNearing B, Brubaker J, Maciel M, Alcala AN, Chakraborty S, Prouty MG, Savarino SJ, Davies BW, Trent MS. 2018. Enterotoxigenic E. Coli virulence gene regulation in human infections. Proceedings of the National Academy of Sciences of the United States of America 115:E8968–E8976.

318. Sheikh A, Charles RC, Sharmeen N, Rollins SM, Harris JB, Bhuiyan MS, Arifuzzaman M, Khanam F, Bukka A, Kalsy A, Porwollik S, Leung DT, Brooks WA, LaRocque RC, Hohmann EL, Cravioto A, Logvinenko T, Calderwood SB, McClelland M, Graham JE, Qadri F, Ryan ET. 2011. In vivo expression of salmonella enterica serotype typhi genes in the blood of patients with typhoid fever in Bangladesh. PLoS Neglected Tropical Diseases 5.

319. Agrawal S, Agrawal XA, Said HM. 2016. Biotin deficiency enhances the inflammatory response of human dendritic cells. Am J Physiol Cell Physiol 311:386–391.

320. Jung HJ, Littmann ER, Seok R, Leiner IM, Taur Y, Peled J, Brink M van den, Ling L, Chen L, Kreiswirth BN, Goodman AL, Pamer EG. 2019. Genome-wide screening for enteric colonization factors in carbapenem-resistant ST258 klebsiella pneumoniae. mBio 10.

321. Scheuerl T, Hopkins M, Nowell RW, Rivett DW, Barraclough TG, Bell T. 2020. Bacterial adaptation is constrained in complex communities. Nature Communications 11.

322. Wetmore KM, Price MN, Waters RJ, Lamson JS, He J, Hoover CA, Blow MJ, Bristow J, Butland G, Arkin AP, Deutschbauer A. 2015. Rapid Quantification of Mutant Fitness in Diverse Bacteria by Sequencing Randomly Bar-Coded Transposons. mBio 6:e00306-15.

323. High molecular weight DNA extraction strategies for long-read sequencing of complex metagenomes https://doi.org/10.1101/2021.03.03.433801.

324. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nature Biotechnology 37:540–546.

325. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PloS one 9:e112963–e112963.

326. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS, Fink I, Pan JN, Yousef M, Fogarty EC, Trigodet F, Watson AR, Esen ÖC, Moore RM, Clayssen Q, Lee MD, Kivenson V, Graham ED, Merrill BD, Karkman A, Blankenberg D, Eppley JM, Sjödin A, Scott JJ, Vázquez-Campos X, McKay LJ, McDaniel EA, Stevens SLR, Anderson RE, Fuessel J, Fernandez-Guerra A, Maignien L, Delmont TO, Willis AD. 2021. Community-led, integrated, reproducible multi-omics with anvi'o. Nature Microbiology 6:3–6.

327. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC bioinformatics 11:119.

328. Lee MD. 2019. GToTree: a user-friendly workflow for phylogenomics. Bioinformatics 35:4162–4164.

329. Eddy SR. 2011. Accelerated profile HMM searches. PLoS Computational Biology 7.

330. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin E v, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov A v, Vasudevan S, Wolf YI, Yin JJ, Natale DA. 2003. The COG database: an updated version includes eukaryotes. BMC bioinformatics2003/09/11. 4:41.

331. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, Ogata H. 2020. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. Bioinformatics (Oxford, England) 36:2251–2252.

332. Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic acids research 28:27–30.

333. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9:357–359.

334. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup 1000 Genome Project Data Processing. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.

335. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. Journal of Molecular Biology 215:403–410.

336. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100.

337. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. Nature biotechnology 29:24–26.

338. Warren DK, Hill HA, Merz LR, Kollef MH, Hayden MK, Fraser VJ, Fridkin SK. 2004. Cycling empirical antimicrobial agents to prevent emergence of antimicrobial-resistant Gram-negative bacteria among intensive care unit patients. Critical care medicine 32:2450–2456.

339. Sarraf-Yazdi S, Sharpe M, Bennett KM, Dotson TL, Anderson DJ, Vaslef SN. 2012. A 9-Year retrospective review of antibiotic cycling in a surgical intensive care unit. The Journal of surgical research 176.

340. Brown EM, Nathwani D. 2005. Antibiotic cycling or rotation: a systematic review of the evidence of efficacy. The Journal of antimicrobial chemotherapy 55:6–9.

341. Burden CJ, Qureshi SE, Wilson SR. 2014. Error estimates for the analysis of differential expression fromRNA-seq count data. PeerJ 2014.

342. Tien M, Fiebig A, Crosson S. 2018. Gene network analysis identifies a central post-transcriptional regulator of cellular stress survival. eLife 7.

343. Singh D, Murashko ON, Lin-Chao S. 2020. Posttranscriptional regulation of tnaa by protein-RNA interaction mediated by ribosomal protein L4 in Escherichia coli. Journal of Bacteriology 202:799–818.

344. Lechner M, Schwarz M, Opitz M, Frey E. 2016. Hierarchical Post-transcriptional Regulation of Colicin E2 Expression in Escherichia coli. PLOS Computational Biology 12:e1005243.

345. Potts AH, Vakulskas CA, Pannuri A, Yakhnin H, Babitzke P, Romeo T. 2017. Global role of the bacterial post-transcriptional regulator CsrA revealed by integrated transcriptomics. Nature Communications 2017 8:1 8:1–15.

346. Ng KM, Aranda-Díaz A, Tropini C, Frankel MR, van Treuren W, O'Laughlin CT, Merrill BD, Yu FB, Pruss KM, Oliveira RA, Higginbottom SK, Neff NF, Fischbach MA, Xavier KB, Sonnenburg JL, Huang KC. 2019. Recovery of the Gut Microbiota after Antibiotics Depends on Host Diet, Community Context, and Environmental Reservoirs. Cell Host & Microbe 26:650-665.e4.

347. Shepherd ES, Deloache WC, Pruss KM, Whitaker WR, Sonnenburg JL. 2018. An exclusive metabolic niche enables strain engraftment in the gut microbiota. Nature 557:434.