EXPERIMENTAL ADAPTATION OF A FREE-LIVING BACTERIUM TO THE ZEBRAFISH DIGESTIVE TRACT

by

JARRETT FARNELL LEBOV

A DISSERTATION

Presented to the Department of Biology
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2019

DISSERTATION APPROVAL PAGE

Student: Jarrett Farnell Lebov

Title: Experimental Adaptation of a Free-Living Bacterium to the Zebrafish Digestive Tract

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Biology by:

Patrick Phillips	Chairperson
Brendan Bohannan	Advisor
William Cresko	Core Member
Karen Guillemin	Core Member
Michael Harms	Institutional Representative

and

Janet Woodruff-Borden	Vice Provost and Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2019

DISSERTATION ABSTRACT

Jarrett Farnell Lebov

Doctor of Philosophy

Department of Biology

June 2019

Title: Experimental Adaptation of a Free-Living Bacterium to the Zebrafish Digestive Tract

Animals have coexisted with an omnipresent and diverse array of bacteria for the entirety of their evolutionary history. As a result, symbioses between animals and bacteria are ubiquitous and can range from mutualism to parasitism. In particular, countless studies have demonstrated the pivotal role that bacteria residing in animal digestive tracts can play in determining animal health and well-being. However, it is still unknown how bacteria evolve the ability colonize animals. Due to the dramatic impacts that animals and bacteria can have on one another's fitness, it is imperative to understand how symbioses between bacteria and their animal hosts originate. Therefore, to elucidate how bacteria evolve novel associations with vertebrate hosts, I serially passaged six replicate populations of a bacterial species with no prior known host associations (*Shewanella oneidensis*) through the digestive tracts of a model vertebrate, zebrafish (*Danio rerio*). After 20 passages through the digestive tracts of groups of larval zebrafish that were derived bacteria free (amounting to approximately 200 bacterial generations), I observed that all six replicate populations evolved to outcompete their unpassaged ancestor in terms of their ability to colonize larval guts. I subsequently sequenced the genomes of four evolved *S. oneidensis* isolates from each replicate population and found

that their competitive advantage stemmed from two distinct classes of mutations that occurred in a mannose sensitive hemagglutinin pilus operon as well as in genes with putative diguanylate cyclase and phosphodiesterase domains. Both types of mutations enhanced bacterial motility, which was associated with increased representation in the aqueous portion of my experimental system and more efficient per capita immigration into zebrafish guts relative to the ancestral *S. oneidensis* reference strain. These increases in motility, were consistent with the behavior of a closely-related *Shewanella* species (*Shewanella sp.* ZOR0012) that has recently been isolated from the zebrafish digestive tract implying that my evolved isolates may be pursuing a similar adaptive trajectory to the one taken by this host-associated species. My results suggest that a non-host-associated microorganism can rapidly improve its ability to colonize hosts, and this study is the first to capture the early adaptive steps necessary to facilitate this transition.

CURRICULUM VITAE

NAME OF AUTHOR: Jarrett Farnell Lebov

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
University of Colorado, Boulder

DEGREES AWARDED:

Doctor of Philosophy, Biology, 2019, University of Oregon
Bachelor of Arts, Ecology and Evolution, 2008, University of Colorado

AREAS OF SPECIAL INTEREST:

Ecology and Evolution

PROFESSIONAL EXPERIENCE:

Research Associate II, Roche – 454 Life Sciences, 2011-2013
Research Associate I, Roche – 454 Life Sciences, 2010-2011
Manufacturing Associate I, Roche – 454 Life Sciences, 2008-2010
Undergraduate Research Assistant, University of Colorado, 2006-2008

GRANTS, AWARDS, AND HONORS:

Genetics Training Grant, Dissertation Stipend, 2015-2018
American Society of Naturalists, Evolution, Student Travel Award, 2017
Gordon Research Seminar, Animal-Microbe Symbioses, Student Travel Award, 2017
William R. Sistrom Memorial Scholarship, 2016-2017
Donald Wimber Fund Award, 2016-2017
Undergraduate Research Opportunity Program, 2006-2008
First place in section in the University of Colorado Engineering Design Expo, 2003

PUBLICATIONS:

Lebov, J. F. and Bohannan, B. J. M. Loss of function mutations in a msh pilus operon increase the ability of a free-living bacterium to colonize a piscine vertebrate host. (In preparation)

Lebov, J. F. and Bohannan, B. J. M. Selection on two separate genetic pathways increases motility and gut colonization in larval zebrafish. (In preparation)

ACKNOWLEDGMENTS

Many people have contributed to the research described in this dissertation. First and foremost, this body of work would not have been possible without the incredible mentorship I received from my dissertation advisor, Brendan Bohannan. Despite having joined his lab with almost no microbiology experience after a turbulent first year of graduate school, Brendan was able to help me find what most interested me and build that into a research program. At key points throughout my doctorate, he provided the guidance I needed to move my project forward, all the while allowing me the autonomy necessary for me to develop into the scientist and problem solver I am today. I cannot thank him enough for the person and teacher he is, and for always causing me to think deeply about the task at hand and the way life works in general.

I would also like to express my sincere gratitude to Karen Guillemin. Karen's lab runs like a finely-tuned machine, and from the start of my dissertation, she made her lab available to me, providing me with an abundance of resources to ensure that my research was successful. Additionally, throughout my dissertation, Karen has had a profound effect on the way I think about host-microbe interactions and what it takes to do mechanistic research in this field. While many people in Karen's lab have been fundamental to my success, I must specifically thank Catherine Robinson and Travis Wiles. Both of these researchers are exceptional microbiologists and they have both played an outsized role in shaping the way I think about microbiology and in helping me plan experiments.

I have also been fortunate to be part of the highly interdisciplinary Microbial Ecology and Theory of Animals community at the University of Oregon, and I would

specifically like to thank Raghuveer Parthasarathy and Tristan Ursell for allowing me to use their microscopes for my research. Furthermore, I want to extend a special thanks to Brandon Schlomann in the Parthasarathy lab for his unbounded intellect, and his patients and tutelage while assisting me with various types of data analysis.

I would also like to thank my Dissertation Advisory Committee (Brendan Bohannan, William Cresko, Karen Guillemin, Michael Harms, and Patrick Phillips) who kept me on track throughout my PhD. Each member of my committee made important contributions to my graduate education. Finally, I must thank the Zebrafish Facility and their associated personnel for ensuring I always had the biological material necessary to conduct my research.

For my parents who always indulged my curiosity, and for my fiancé who keeps my life in balance.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

Figure Page

LIST OF TABLES

Table Page

CHAPTER I:

INTRODUCTION

On Earth, because bacteria are ubiquitous and vastly outnumber their potential animal hosts, animals live under constant exposure to bacterial life. Consequently, animals have established persistent symbiotic relationships with an array of bacterial species, and these encounters can have dramatic impacts on the reproductive success, or fitness, of both animals and the bacteria they encounter. For example, some bacteria can behave as pathogens that make animals ill, however because the entirety of animal evolution has taken place in the presence of bacteria (Hug et al., 2016), exposure to bacteria is also critical to proper animal health and development (Dominguez-Bello et al., 2019; Luczynski et al., 2016; Bordenstein & Theis, 2015; Sommer and Bäckhed 2013; Flint et al., 2012). Similarly, although animals provide bacteria with a consistent source of nutrients, as many bacterial species aid in animal digestion (Flint et al., 2012), animal hosts also present bacteria with a plethora of selective forces that can eradicate them or prevent their colonization in the first place (Bakke et al., 2015; Ley et al., 2008; Donaldson et al., 2016; Sass et al., 2010; McLoughlin et al., 2016; Quinn et al., 2018; Friedman et al., 2018; Ottman et al., 2017; Segal, 2005, Weiss & Schaible, 2015; Islam et al., 2011). As a result, host-associated communities of microorganisms do not appear to be composed of a random sample of the bacteria a host encounters, but instead hosts tend to assemble communities that are distinct from those found in their surrounding environment (Bakke et al., 2015; Ley et al., 2008). To date, work on host-microbe interactions has been performed across a set of distantly-related host systems, each harboring their own unique community of microorganisms (Mazel et al., 2018; Sachs et

al., 2011, Rawls et al., 2006). This suggests that free-living bacteria, or those species that are not routinely found in close physical associations with hosts, have repeatedly evolved novel symbioses with animals throughout evolutionary history. Further, this body of work focuses almost exclusively on bacterial species that have pre-established relationships with animal hosts, and it is not currently known how free-living bacteria evolve host-associations *de novo*. Given the impacts that bacteria can have on animal health, it is imperative to understand how they transition from a free-living state to a host-associated one.

Host-associated bacteria, which are those species that are commonly isolated from animal hosts, rely on a myriad of traits to colonize animals (Sarkar et al., 2018; Hsiao et al.,2006; Raina et al., 2019; Ribet & Cossart, 2015; Lee et al., 2013), and an abundance of research has established the gut as a primary site of symbiotic interaction. In particular, studies have observed dramatic differences in composition of the gut microbiota of healthy and unhealthy individuals (Hsiao et al., 2013; Morgan et al., 2012; Ley et al., 2005). Additionally, the gut microbiota has been shown to alter gene expression in the gut (Fukushima et al., 2003, Larsson et al., 2011; Cash et al., 2006) and effect tissue maturation in host digestive tracts, as well as elsewhere in the body (Bouskra et al., 2008; Bates et al., 2006; Broderick et al. 2014; Heijtz et al., 2011; Rawls et al., 2004). Even more convincingly, fecal microbiota transplants, in which gut microbial communities are transferred from one host to another, have suggested a causative role of the microbiota in reducing the incidence and severity of disease (Jacob et al., 2006; Van Nood et al., 2013), modulating behavior (De Palma et al., 2017; Bercik et al., 2011), and in altering host metabolic function (Turnbaugh et al., 2006). In some cases, even administering cultured

isolates of single bacterial species can mimic some of these same effects (Hsaio et al., 2013; Buffie et al., 2015). Collectively, this research underscores the outsized influence that members of the gut microbiota can have on host fitness.

Given the central role the gut plays as a site of host-microbe interaction, it is important to consider how bacteria might evolve to gain access specifically to the gut. Because animals live in a world teeming with bacteria, animal hosts can acquire new microbes from a range of sources. One of the most high-fidelity modes of microbial acquisition is known as vertical transmission, where microbes transmit transgenerationally from parent to offspring (Dominguez-Bello et al., 2010). This mode of transmission virtually ensures that offspring will be exposed to a microbial symbiont, thus maximizing the symbiont's potential to re-establish itself as part of the new generation's microbiota. However, human studies have suggested that only a small fraction of the 100's of species found in the gut are heritable (Goodrich et al., 2016; Rothschild et al., 2018), and vertical transmission requires that an inherited microbe already has an association with the parent. Therefore, studying this mode of transmission to understand how bacteria initiate symbiotic relationships with hosts may miss important strategies employed by hosts and microbes which make symbioses more likely. A more fruitful mode of transmission to study to understand the origin of host-microbe associations, would be horizontal transmission, whereby a host's microbiota is acquired from exposure to different components of the environment as they develop, rather than from physical contact with a parent as a result of an organism's normal life cycle (Bright & Bulgheresi, 2010).

From a bacterial perspective, horizontal transmission requires that a bacterial lineage must be able to survive outside a host long enough to encounter, and then colonize a host. Merely encountering a host however, provides no guarantee of reproductive success for bacteria. While hosts can serve as a potential buffer from dramatic temperature swings, and provide shelter from harmful solar radiation, as well as a steady flow of nutrients in the form of partially digested food, detritus, or secretions, they can also serve as an adversary to potential colonists. Host immune systems can target and eliminate unrecognized microbial invaders (McGuinness et al., 2016; Parsons et al., 2019). Additionally, many host digestive tracts contain acidic conditions and bile that can make it difficult for bacteria to survive (Wu et al., 2014; Duncan et al., 2009; Islam et al., 2011; Jones et al., 2008). Finally, some forms of potential nutrients, such as host-secreted glycans, may favor the growth of well-adapted microbial species at the expense of others (Schluter et al., 2012; Koropatkin et al., 2012), which could result in the redeposition of a maladapted lineage back to the external environment. The more time an organism spends in the external environment, the more selective pressures associated with a free-living lifestyle will shape a lineage's evolutionary trajectory. Because of these dynamics, it is unsurprising that bacterial communities found in association with vertebrate hosts are compositionally distinct from those sampled from non-vertebrate host environmental sources (Bakke et al., 2015; Ley et al., 2008).

Still, encounters between vertebrates and bacteria happen on a constant basis. While some of these encounters may result in eventual bacterial death due to a fitness cost imparted by the host, others may result in the continued propagation of a bacterial species. Even if these host-microbe interactions are transient in nature, surviving lineages

will have undergone a period of selection, however brief, that may better prepare them for subsequent encounters with a similar host. As these lineages continue to evolve, genotypes could arise that encode traits which may either prolong a given cell line's interaction with a future host or may increase the frequency with which that cell line has future encounters with a host. Each encounter with a subsequent host would again serve as a selective filter that enriches for genotypes capable of survival in a host. Over time if host-surviving lineages were not outcompeted while in non-host environments, and this cycle continued with great enough frequency, a bacterial species could improve its association with a host as evolution eventually produced evermore host-adapted genotypes. In this way, repeated host exposure could select host-associated lineages over evolutionary time.

This conceptual framework, while consistent with current evolutionary theory, obviously relies heavily on a multitude of chance events, each of which could result in a multitude of diverse outcomes. Many of these outcomes may lead to bacteria that are better-suited for a free-living existence than a host-associated one. Of the minority of outcomes that result in host-association, many questions remain. Namely, which traits improve host colonization? At what level (gene sequence, gene product, physiology, behavior, or population) do these traits operate? How does the architecture of a genome — a genome's sequence and physical structure — bias evolutionary events that might lead to host-association? How reproducible are the evolutionary events that lead to host-association, both in terms of the actual events that occur, and the order in which those events happen? Which aspect of the host environment imparts the strongest selection pressures? For a given bacterial lineage, it is crucial to address these questions if we are

to understand both the likelihood that a bacterium could become host-associated, and the process such a bacterium might follow to become host-associated.

To confront these unknowns, I have sought to understand how a bacterium might evolve a novel host-microbe symbiosis by serially passaging a bacterial species, with no documented history of an association with metazoan hosts (*Shewanella oneidensis*), through the digestive tract of a model vertebrate, zebrafish (*Danio rerio*). The zebrafish system is an ideal system in which to address these questions. The zebrafish gut, like other vertebrates, is composed of a diverse set of selective pressures including various mucosal surfaces (Taormina et al., 2017), the presence of both innate and adaptive immune systems (Rolig et al., 2015; Stagaman et al., 2017), multiple potential nutrient sources (Semova et al., 2012; Wong et al., 2015; Schluter & Foster, 2012), strong chemical gradients (Flores et al., 2010), and frequent physical disturbances (Burns & Guillemin, 2017; Wiles et al., 2016). Additionally, zebrafish contain complex gut microbial communities whose constituents show niche preferences within the digestive tract (Schlomann et al., 2018; Wiles et al., 2016), and the zebrafish gut microbiota shows the ability to influence host health and development in ways that are common among vertebrates (Burns & Guillemin, 2017, Rawls et al., 2004). Further, much is known about zebrafish life-history, including the developmental stage when larvae are susceptible to bacterial colonization, and there are established protocols for the dissection of larval digestive tracts so that I can specifically study gut-associated bacteria (Burns and Guillemin, 2017). Zebrafish are also highly fecund and easy to cohouse (Burns & Guillemin, 2017), which makes it easy to investigate how bacteria adapt to groups of social hosts rather than an individual host (Suriyampola et al., 2016). Lastly, the process

of colonization can be studied simply by inoculating the aqueous environment that zebrafish inhabit with potential bacterial colonists (Robinson et al., 2018). Together, these unique features facilitate the examination of a large fraction of the parameters that might support novel host adaptation.

Given the enormous diversity of bacteria on Earth however, I also needed to carefully consider which bacterial species might maximize my ability to understand how bacteria initiate host-microbe symbioses. *S. oneidensis* (MR-1; manganese reducer 1, so named after the observation *S. oneidensis* could reduce manganese; Meyers & Nealson, 1988; Venkateswaran et al., 1999) poses several advantages that make it an excellent candidate in which to study the evolution of *de novo* host adaptation. Its gene functions are well-annotated (Deutschbauer et al., 2011; Rodrigues et al., 2011; Ong et al., 2014; Heidelberg et al., 2002), it is amenable to genetic manipulation (Saville et al., 2010; Thormann et al., 2006), and much is known about its physiology, metabolism, and how it responds to various types of stress (Li et al., 2010; Liu et al., 2005; Saville et al., 2010; Bouhenni et al., 2010; Chao et al., 2013; Jiang et al., 2014; Gao et al., 2004; McLean et al., 2008). Additionally, the *Shewanella* genus is commonly observed in the digestive tracts of larval zebrafish (Stephens et al., 2016), although importantly, MR-1 in particular has never been isolated from a eukaryotic host. Despite this, MR-1 can colonize the guts of bacteria-free (BF) larval zebrafish (Figure 3A), in the absence of competitors, when it is added to their aqueous environment, which allows for the interrogation of the factors that support this process. Finally, compared to another *Shewanella* species isolated from the larval zebrafish digestive tract with which MR-1 shares a recent common ancestry (Figure 2, A and B), MR-1 neither achieves the same abundances in the zebrafish

digestive tract as this zebrafish-associated species (Figure 3A), nor is it able to compete effectively with this species to colonize the digestive tract (Figure 3B). Together these features imply that MR-1 has ample room to improve its symbiotic affinity with this novel host.

The goal of the research described herein is to determine how non-host-associated, or free-living, bacteria evolve to initiate transitions towards host-association. In Chapter II, I will outline a selection regime capable of increasing the ability of a free-living bacterial species, MR-1, to colonize the digestive tracts of BF larval zebrafish. Further, Chapter II will show evidence that establishes the genetic basis undergirding the ability of host-adapted MR-1 strains to outcompete their free-living ancestor and identify the physiological effects of adaptive mutations in an evolved MR-1 isolate. I will then dissect out which components of the experimental environment provided niche space into which evolved MR-1 lineages could expand. Finally, I will conclude this chapter by identifying behaviors which could plausibly explain the expansion of evolved lineages into new niches and facilitate host colonization by a formerly naive bacterial species. Chapter III will describe the consistency of selection imparted by the serial passage scheme outlined in Chapter II. In doing so, it will compare the fitness and phenotypes of two evolved bacterial isolates with distinct sets of adaptive mutations that result in improved colonization of the zebrafish gut. Additionally, this chapter will consider whether evolved MR-1 strains are following a convergent adaptive trajectory with respect to a closely-related *Shewanella* species that has been isolated from the zebrafish gut. Ultimately, I will conclude in Chapter IV with a synthesis of my findings and a discussion of how they fit within the broader context of host-microbe research.

CHAPTER II:

LOSS OF FUNCTION MUTATIONS IN AN MSH PILUS OPERON INCREASE THE ABILITY OF A FREE-LIVING BACTERIUM TO COLONIZE A PISCINE VERTEBRATE HOST

Introduction:

Given the outsized role that gut bacteria can play in influencing the health of their animal hosts, many studies have focused on trying to understand which traits maintain existing host-microbe symbioses. To accomplish this, researchers often analyze the genomic content (Licandro-Seraut et al., 2014) and expression patterns (List et al., 2018; Rey et al., 2010; Xu et al., 2003; Fabich et al., 2011) of bacterial taxa that are commonly found in host guts. In these studies, variability in genomic content can be vital to helping researchers identify which genomic features govern host-microbe interactions. This variability can be generated *de novo* using various mutagenesis schemes (Goodman et al., 2009; Lee et al., 2013; Stephens et al., 2015), or it can be culled from a large set of isolates from unique sources (Lee et al., 2017; Yahara et al., 2017; Arredondo-Alonso et al., 2019, Frese et al., 2011). In either case, comparative genomics approaches, in which sets of genomes are aligned and then examined for differences, can be implemented to identify candidate loci that are important for host-association. Such analyses can then be followed up with genetic manipulations to determine whether candidate loci indeed contribute to host-microbe interactions (Lee et al., 2013; Mazmanian et al., 2005, Leatham et al., 2005; Hsiao et al., 2006). While tremendously informative to our understanding of the host-microbe conversation, lists of adaptive candidates can be too long to interrogate individually, many bacteria are not amenable to genetic manipulation, and gene annotations are limited in their ability to explain how candidate genes might

function under the diverse sets of conditions that are found within hosts. Further, these approaches rely on mere snapshots of preexisting biological relationships that have already evolved, and they are not capable of identifying the full range of adaptations which engender and maintain host-microbe symbioses.

A more comprehensive understanding of the breadth of traits that facilitate host associations can be obtained by analyzing how bacteria evolve and respond to their hosts over time. To see why this is, it is important to consider some of the processes that govern bacterial evolution. Because bacteria predominantly reproduce asexually, bacterial populations commonly undergo clonal interference, whereby genotypes within a population compete for representation (Barrick & Lennski, 2013; Fogle et al., 2008, de Visser & Rozen, 2006; Wilke, 2004, Gerrish & Lenski, 1998). The effect of this phenomenon is that bacteria whose genomes house the most beneficial mutations tend to increase their frequency within a population at the expense of other competitors. In some cases, this means that even lineages with beneficial alleles can go extinct if they are competing against other lineages with alleles of higher adaptive value (Fogle et al., 2008; Rozen et al. 2002, Wilke, 2004; Barroso-Batista et al., 2014). In these cases, extinct lineages would leave no trace of the beneficial traits they might have had, limiting our understanding of the full breadth of traits that are important for survival within a given environment.

Additionally, in many cases the advantage imparted by adaptive alleles can be highly dependent on genomic context and epistatic interactions with other loci (Blount et al., 2008, Khan et al., 2011; Chiotti et al., 2014). During evolution, genomic content and architecture can be altered (genes can be inserted, deleted, or their relative position in the

genome can be rearranged). Because some information may be lost as organisms evolve, this makes it difficult to interpret how genes depend on each other to encode specific functions by only considering the genomic content of extant species. For example, changes in genomic content could potentiate the evolution of a new trait. This trait, once evolved, could then develop new dependencies on other genes which could make the trait's old dependencies obsolete. The alleles responsible for the initiation of this new trait could then be recombined out of the genome over time, or they could revert to their original state (Bull et al., 1997), leaving no trace that they had the potential to modulate this derived trait in the organism being examined, or in other organisms with similar gene networks. By observing how an organism's genome changes over time as it adapts to a new environment, it is possible to gain a more complete understanding of all the genes involved in a given function, and how changes in gene networks might alter that function to impart adaptation.

Many researchers have attempted to circumvent these limitations by utilizing an approach known as experimental evolution in which an organism is passaged through an environment to study how it adapts to that environment. When combined with the periodic genomic sequencing of evolving lineages, this strategy allows for the observation of evolutionary events right after they happen. Phenotypic and fitness assays can then be performed, and this information can be synthesized to understand how evolved mutations improve fitness. In this way, snapshots from multiple time points can be integrated to paint a detailed picture of the full set of traits that are important to host-association, how those traits depend on each other, and in what order those traits are likely to evolve (Barroso-Batista et al., 2014; Khan et al., 2011, Long et al., 2015, Mani

& Clarke, 1990). Given their propensity for short generation times, many researchers have relied on experimental evolution to uncover how various bacterial species adapt to a range of *in vitro* environments (Rainey & Travisano, 1998; Wiser et al., 2013; Traverse et al., 2013). Recently, the field of experimental evolution has expanded to include some host environments. These studies mostly rely on the adaptation of bacterial species to single hosts which are housed separately (Giraud et al., 2001; Giraud et al., 2008; Barroso-Batista et al., 2014, Leatham et al., 2005), and the methods used in some studies often artificially bypasses some host filters (intraperitoneal injection, gavage, etc.; Barroso-Batista et al., 2014, Nilsson et al., 2004). In nature however, hosts interact with their environment as well as with each other. Therefore, these studies do not address how host adaptation might be affected by ecological considerations such as host-environment or host-host interactions. Further, as these studies show, previous attempts at exploring experimental evolution in hosts have utilized bacterial symbionts that have been isolated from hosts, and it is not known how these relationships originated.

In recognition of the fact that animals are in regular contact with an array of free-living bacteria, and that bacteria can have large impacts on animal fitness, I sought to understand how a free-living bacteria might improve its ability to colonize an animal host by tracking the evolution of a bacterium with no prior known host-associations (*Shewanella oneidensis*), as it initiates an association with cohorts of a model vertebrate, zebrafish (*Danio rerio*). To accomplish this, I serially passaged six replicate MR-1 populations through the digestive tracts of groups of BF larval zebrafish. Each passage involved the inoculation of flasks containing 10-15 BF larval zebrafish at four days post fertilization (dpf) with MR-1 populations at ~10^3 CFU/mL. MR-1 populations were then

incubated with larvae for 72 hours during which time the bacterial populations grew to ~10^6 CFU/mL and colonized the larval guts. I then dissected and homogenized 10 larval digestive tracts and used the homogenate to inoculate a subsequent batch of BF larvae (Figure 1). After each bout of dissections, a sample of the homogenized guts was archived in a freezer stock that was stored at -80 °C. I repeated this cycle for 20 passages, at which point I assessed the competitive fitness of isolates from each replicate population relative to their unpassaged ancestor. I found that evolved MR-1 isolates were selected to outcompete their ancestor in the larvae-conditioned aqueous environment outside the larvae, as well as translocate into larvae more efficiently, with no apparent increase in components of fitness related to growth inside larvae. Further, I found strong evidence for selection on a pilus operon which may underlie these enhanced attributes. In this way, formerly free-living MR-1 populations maximized their association with a host digestive tract indirectly by adapting to external aspects of the host environment, and potentially altering their physiology to more efficiently bypass host filters.

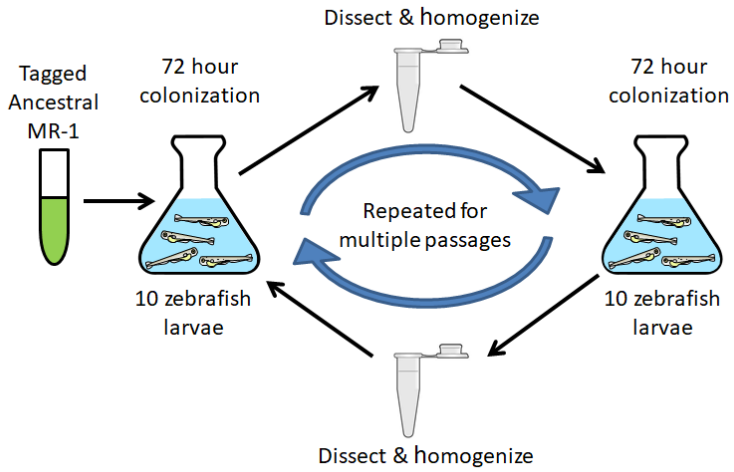

Figure 1: Serial passage scheme. Serial passage was started with a tagged ancestral culture of *MR-1* which was used to inoculate a group of 4 days post fertilization (dpf) bacteria free (BF) larvae. BF larvae were incubated with *MR-1* populations for 72 hours, at which point 10 larval guts were dissected and homogenized. A sample of the homogenate was then used to create a freezer stock held at -80 °C, and the rest of the homogenized guts were used to inoculate a subsequent set of 4 dpf BF larvae. This cycle was repeated for 20 passages.

Methods:

Zebrafish husbandry:

To ensure animal specimens were treated ethically in all experiments involving zebrafish, I adhered to the standard protocols and procedures approved by the University of Oregon Institutional Animal Care and Use Committee (IACUC protocol: 15-98). BF derivations were carried out as described in Melancon et al., 2017. Details about larval gut dissections can be found the Serial passage section.

Bacterial strains:

My ancestral reference *S. oneidensis* (MR-1) and Shew-Z12 strains were obtained from Karen Guillemin's laboratory at the University of Oregon. All modifications to MR-1, including Tn7-mediated gfp and dTomato insertions as well as allelic exchange applications, were carried out as described in Wiles et al., 2018. Whole gene deletion allelic constructs were created via a splice by overlap extension protocol (Wiles et al., 2018), whereas evolved mutation allelic constructs were created via amplification of mutated segments of evolved genomes. All *S. oneidensis* strains were cultured in TSB at 30 °C under shaking conditions unless otherwise specified.

Serial passage:

My experimental evolution serial passage scheme was similar to the one outlined by Robinson et al. 2018. 5 mL overnight tryptic soy broth (TSB) cultures of MR-1 isolates tagged with either green fluorescent protein (MR-1gfp) or dTomato fluorescent protein (MR-1dT) were diluted 1:100 in TSB and allowed to grow out to late log phase (4-5 hours). Six replicate ancestral populations were then generated by combining subcultures of MR-1gfp and MR-1dT at a 1:1 ratio. These mixtures allowed us to infer the occurrence adaptive events based on fluorescent tag frequency changes observed throughout the experiment. Beneficial mutations occurring in a tagged genomic background should cause the frequency of that tag to increase over time. Next, 10 μL of each of these replicate ancestral populations were used to inoculate larval flasks containing ~15 mL of EM and ~15 BF larval zebrafish at 4 days post fertilization (dpf) (inoculating MR-1 densities were ~10^6 CFU/mL). Larvae were then incubated with MR-1 populations at 28

˚C for 72 hours. At 7 dpf, 10 larvae were euthanized with tricane, mounted on a glass slide, and their digestive tracts were dissected out. Glass slides were coated with 3 % methylcellulose to help immobilize larvae during dissections. Following dissections, all 10 digestive tracts culled from each flask were placed in a single 1.7 mL tube containing 500 μL EM and ~100 μL 0.5 mm zirconium oxide beads (Next Advance, Averill Park, NY, US). The contents the larval guts in each these tubes were then immediately homogenized using a bullet blender tissue homogenizer (Next Advance, Averill Park, NY, US) for 60 seconds at power 4. To preserve my ability to revive replicate populations after each passage, I created freezer stocks by using a pipette to mix 200 μL from each homogenized tube with 200 μL of 50 % glycerol (25 % glycerol final concentration). These freezer stocks were stored at -80 °C. The remaining contents of each homogenized tube was then stored at 4 ˚C for 0-14 days, at which point ~250 μL were sampled to inoculate a subsequent set of BF larval flasks (~15 larvae in ~15 mL EM). Upon inoculation, 100 μL of each larval flask was dilution-plated in triplicate to quantify the inoculating population densities (typically around 10^3 CFU/mL) and determine tag frequencies. This cycle was repeated for 20 passages. All six replicate evolving populations were maintained separately throughout my experiment.

Comparative genomics:

I submitted my MR-1 and Shew-Z12 strains to the Washington State University Molecular Biology and Genomics Core (WSUGC) for long read sequencing. Genome assembly for MR-1 was conducted by WSUGC, whereas genome assembly for Shew-Z12 was conducted in house with Canu v. 1.7.1 (Koren et al., 2017). To generate

annotation files for these genomes, I relied on Prokka v1.12 (Seemann, 2014), and RAST v. 2.0 (Aziz et al., 2008).

Phylogenetics:

Using Integrated Microbial Genomes and Microbiomes (IMG/M, https://img.jgi.doe.gov/; Chen et al., 2018), I collated a set of 16S ribosomal RNA genes from 28 *Shewanella* species, 2 *Vibrio* species, and 1 *Aeromonas* species (See Appendix Table 2 for metadata). These 16S rRNA genes were entered into Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) to generate a multiple sequence alignment file and a subsequent Newick-formatted phylogenetic tree file. This file was then visualized with FigTree v1.4.4 (Rambaut & Drummond, 2012).

Genome comparisons between *S. oneidensis* and non-*S. oneidensis* species:

Average sequence identity (ANI) was calculated using the EZBioCloud online ANI calculator (Yoon et al., 2015) to quantify the ANI between *S. oneidensis* and Shew-Z12. Whole genome amino acid sequence alignment visualizations were conducted using the Seed Viewer v. 2.0 sequence comparison tool (Overbeek et al. 2013).

Specific gene and operon comparisons between *S. oneidensis* and Shew-Z12:

ANI was calculated using the same tool described in my whole genome comparisons above. For mshOP comparisons, I separately concatenated the amino acid sequences of each gene in the mshOP of *S. oneidensis* and Shew-Z12 in series by relying on my RAST-annotated files. I then generated multiple sequence alignment (msa) files using Clustal Omega web tool (Sievers et al., 2011) that compared these mshOP sequences and used them to depict sites of divergence along the mshOP (Figure 5). Visualizations of single gene or multi-gene comparisons between *S. oneidensis* and Shew-Z12 were created

using Clustal Omega-based msa files that were imported into Jalview2 (Waterhouse et al., 2009). To highlight regions of divergence within genes I color-coded my comparisons using the color by annotation feature of Jalview2 (Figure 5; Figure 17). This feature color-codes amino acid comparisons per site based on biochemical conservation.

Evolved mutation calling:

I obtained 4 isolates per evolved replicate population (24 isolates total) by using an inoculation loop to dilution streak a sample from each population's freezer stock on TSA plates (one plate per evolved population, totaling 6 plates) and incubating plates at 30 °C for 24 hours. For each population, I overnight cultured four colonies that resulted after 24 hours of growth, and then created freezer stocks (stored at -80 °C) that consisted of a 1:1 mixture of each cultured isolate and an equal volume of 50 % glycerol (final concentration: 25 % glycerol). To create genomic libraries for each evolved isolate, I used an inoculation loop to generate overnight cultures from each isolate's corresponding freezer stock, and then extracted genomic DNA from each culture using the Promega Wizard® Genomic DNA Purification Kit (Catalog #: A1120). Single-end 150 bp libraries were generated from these genomic DNA extractions according to the Nextera XT DNA Library Prep Kit Reference Guide (Document #: 15031942 v02), and these libraries were sequenced on the Illumina HiSeq 4000. Following this same protocol, I also sequenced the genomes of my ancestral MR-1gfp and MR-1dT strains on the Illumina HiSeq 4000. To identify candidate adaptive mutations, I relied on breseq 0.31.0 (in consensus mode) to compare each Illumina sequenced isolate to the annotated ancestral reference separately, and then looked for single nucleotide polymorphisms (SNPs) and indels that were present in each evolved isolate, but absent in MR-1gfp and MR-1dT isolates. With

the exception of mshOP genes, the gene annotations for the mutations listed in Table 1 were determined by Prokka v1.12 (Seeman, 2014). The mshOP gene annotations were determined by RAST v 2.0.

Whole experimental system competitions:

5 mL overnight TSB cultures of competing strains were diluted 1:100 in TSB and allowed to grow out to late log stage (4-5 hours). 500 μL of each competitor was then mixed in a single 1.7 mL tube so that competitors were at an approximate 1:1 ratio. Competition mixtures were pelleted (7000 rcf for 5 min) and resuspended in 1 mL sterile EM. Resuspended competition mixtures were diluted 1:100, and 7.5 μL of these dilutions were used to inoculate BF larval flasks containing ~15 mL of EM and ~15 larvae at 4 dpf. Immediately following inoculation, triplicate 100 μL samples from each competition flask were dilution plated to establish the inoculation ratio of competitor1:competitor2 (CFU/mL). After inoculation, I incubated larval flasks at 28 ˚C for 72 hours. At 7 dpf, multiple larvae per flask were euthanized with tricane, and their guts were dissected and individually placed in 1.7 mL tubes containing 500 μL of sterile EM and ~100 μL 0.5 mm zirconium oxide beads (Next Advance, Averill Park, NY, US). The contents the larval guts in each these tubes were then immediately homogenized using a bullet blender tissue homogenizer (Next Advance, Averill Park, NY, US) for 60 seconds at power 4. Homogenized tubes were then dilution plated to determine the CFU/gut for each competitor. A competitive index (CI) was calculated for each dissected gut by dividing the ratio of competitor1:competitor2 found in each gut by the mean inoculation ratio

determined from the triplicate measurements for each corresponding flask, CI = $\left(\frac{Competitor1{:}Competitor2_{guts}}{Competitor1{:}Competitor2_{inoculum}}\right)$.

Rich media competitions:

Overnight TSB cultures of competing strains were diluted 1:100 in TSB and allowed to grow out to late log stage (4-5 hours). 500 μL of each competitor was then mixed in a single 1.7 mL tube so that competitors were at an approximate 1:1 ratio. Competition mixtures were pelleted (7000 rcf for 5 min) and resuspended in 1 mL TSB. Resuspended competition mixtures were diluted 1:100, and 5 μL of these dilutions were added to 10 mL TSB in a 20 mL test tube. Immediately following inoculation, triplicate 100 μL samples from each competition culture tube were dilution plated to establish the inoculation ratio of competitor1:competitor2 (CFU/mL). I incubated each competition at 30 ˚C for 24 hours with agitation (angled back and forth rocker, 60 rpm), at which point I again took triplicate 100 μL samples from each competition tube and dilution plated them to quantify the CFU/mL for each competitor. Competitive indices were calculated by dividing the final CFU ratio of competitor1:competitor2 by the inoculation ratio, CI = $\left(\frac{Competitor1{:}Competitor2_{final}}{Competitor1{:}Competitor2_{inoculum}}\right)$.

Transmission electron microscopy (TEM):

Samples were prepped for TEM imaging as described in Jones et al. 2015. Briefly, freezer stocks of my ancestral reference *S. oneidensis* strain were revived in a 5 mL TSB overnight culture with shaking at 30 °C. The next day, the culture was diluted 1:100 in

TSB and gown under the same conditions until an optical density (OD) of 0.3-0.5 was reached (2-3 hours). Subcultures were diluted to an OD of 0.1-0.2, and then 2 μL were loaded onto a 300-mesh carbon-coated Formvar grid (Ted Pella product #: 01820). After a 5-minute drying period, the gridded sample was washed twice with deionized water. Water washes were accomplished by transferring the loaded grid to successive 5 μL beads of water on parafilm for 30 seconds at a time. Following the water washes, I stained my sample with a 2 % (w/v) solution of uranyl acetate for 90 seconds and then placed the stained grid on a 5 μL bead of water on parafilm for 30 seconds. Imaging was conducted with a FEI Tecnai G2 Spirit TEM STEM.

Biofilm assay:

Triplicate biological replicate overnight TSB cultures of strains of interest were diluted 1:100 in TSB and allowed to grow out to mid-late log stage (3-4 hours). Optical densities were measured at 600 nm for each overnight culture. 1 mL of subcultured strains were pelleted (7000 rcf for 5 min), and the pellets were resuspended with a volume of sterile EM that would correspond with a TSB optical density of 1.0 if read at 600 nm. 150 μL of each resuspension were added to four wells of a round-bottom 96 well polystyrene plate (Greiner Bio-One, catalogue #: 650185) per biological replicate. The plate was incubated at 30 ˚C for 24 hours, and the volume of each well was removed with a multichannel pipette. The wells were rinsed twice with 200 μL of sterile EM, and each well was stained with 180 μL 0.1 % crystal violet (CV). The plate was incubated at room temperature for 10 min, at which point the crystal violet was removed with a multichannel pipette, and the wells were again rinsed twice with 200 μL of sterile EM.

The CV was solubilized with 100 % dimethyl sulfoxide (DMSO) for 15 min with agitation (~180 rpm on a rotating minishaker), and 100 μL of the solubilized CV was added to 100 μL of 100 % DMSO in a flat-bottom 96 well polystyrene plate (Corning Incorporated, reference #: 3595). Optical densities were then measured for each well at 570 nm.

Colonization level assessments:

5 mL overnight TSB cultures of strains of interest were diluted 1:100 in TSB and allowed to grow out to late log stage (4-5 hours). 1 mL of subcultured strains were then pelleted (7000 rcf for 5 min), and the pellets were resuspended in 1 mL sterile EM. This resuspension was then diluted 1:100, and 7.5 μL of this dilution was used to inoculate BF larval flasks containing ~15 mL of EM and ~15 larvae at 4 dpf. After inoculation, I incubated larval flasks at 28 ˚C for 72 hours. At 7 dpf, multiple larvae per flask were euthanized with tricane, and their guts were dissected and individually placed in 1.7 mL tubes containing 500 μL of sterile EM and ~100 μL 0.5 mm zirconium oxide beads (Next Advance, Averill Park, NY, US). The contents the larval guts in each these tubes were then immediately homogenized using a bullet blender tissue homogenizer (Next Advance, Averill Park, NY, US) for 60 seconds at power 4. Homogenized tubes were dilution plated to determine the CFU/gut.

Colonization level over time:

Overnight TSB cultures of competing strains were diluted 1:100 in TSB and allowed to grow out to late log stage (4-5 hours). 500 μL of each competitor was then mixed in a

single 1.7 mL tube so that competitors were at an approximate 1:1 ratio. Competition mixtures were pelleted (7000 rcf for 5 min) and resuspended in 1 mL sterile EM. Resuspended competition mixtures were diluted 1:100, and 7.5 μL of these dilutions were used to inoculate BF larval flasks containing ~30 mL of EM and ~30-35 larvae at 4 dpf. At 0, 2, 4, and 8 hours post inoculation, 100 μL of the EM containing the larvae was sampled and dilution plated to determine the CFU/mL for each competing strain, and then five larval guts were immediately dissected, and their contents dilution plated (described in *Whole experimental system competitions* above), to determine the CFU/gut of each competing strain.

Larvae-conditioned media (LCM) competitions:

Overnight TSB cultures of competing strains were diluted 1:100 in TSB and allowed to grow out to late log stage (4-5 hours). 500 μL of each competitor was then mixed in a single 1.7 mL tube so that competitors were at an approximate 1:1 ratio. Competition mixtures were pelleted (7000 rcf for 5 min) and resuspended in 1 mL sterile EM. Resuspended competition mixtures were diluted 1:100, and 7.5 μL of these dilutions were used to inoculate BF larval flasks where larvae had been removed at 4 dpf. Each of these BF larval flasks had contained ~15 mL of EM and ~15 larvae that had been conditioning their EM over the course of four days. Multiple BF larval flasks had been combined into a single sterile vessel, and the larvae were removed using a 1.5 mL bulb pipette, leaving behind only the LCM. ~15 mL of LCM was then aliquoted into fresh sterile polystyrene culture flasks. These flasks were inoculated as stated above and triplicate 100 μL LCM samples were taken per competition flask and dilution plated on TSA to establish an

initial competition ratio (competitor1: competitor2). Each competition flask was incubated at 28 ˚C for three days (to mimic a single passage during my serial passage procedure) at which point 100 μL samples were again taken in triplicate and dilution plated to establish a final competition ratio (competitor1:competitor2). Competitive indices were calculated by dividing the final CFU ratio of competitor1:competitor2 by the inoculation ratio, $\mathrm{CI} = \left(\frac{Competitor1{:}Competitor2_{guts}}{Competitor1{:}Competitor2_{inoculum}}\right)$.

LCM growth dynamics:

Overnight TSB cultures of strains of interest were diluted 1:100 in TSB and allowed to grow out to late log stage (4-5 hours). 1 mL of subcultured strains were then pelleted (7000 rcf for 5 min), and the pellets were resuspended in 1 mL sterile EM. Resuspensions for each strain were then diluted 1:100, and 7.5 μL of each dilution was used to inoculate separate ~15 mL LCM flasks (created as described in LCM competitions assays above). 100 μL samples from LCM flasks were then dilution plated on TSA every few hours to assess population dynamics over time by quantifying the CFU/mL at each time point.

Immigration and in vivo growth assays:

Separate 5 mL overnight TSB cultures of competitor strains of interest were diluted 1:100 in TSB and allowed to grow out to late log stage (4-5 hours). 1 mL of subcultured strains were then pelleted (7000 rcf for 5 min), and the pellets were resuspended in 1 mL sterile EM. Resuspensions for each competitor were then diluted 1:100, and 7.5 μL of each dilution was used to inoculate separate ~15 mL LCM flasks (created as described in LCM

competitions assays above). Each competitor was incubated in its own LCM flask at 28 ˚C for 12-15 hours at which point two flasks (one containing each competitor) were combined into a single polystyrene petri dish such that the petri dish contained a ~30 mL competition mixture. 30-35 4 dpf larvae were then added to each competition mixture dish, and 100 μL samples were immediately taken in triplicate from each dish. These samples were dilution plated to establish an inoculating competition mixture (competitor1:competitor2). After a 40-60 min incubation period at 28 ˚C, the LCM was removed and replaced with 100 mL sterile EM for a total of three rinses. These media exchanges were conducted to reduce the density of each *S. oneidensis* competitor in the EM so that both populations were below a threshold where colonization would readily occur (colonization threshold > 10^4 CFU/mL; Figure 8A & 8B). 10 larval guts were then dissected as described above in my whole experimental system competitions. By dilution plating the contents of each dissected gut, I established a competition ratio for each gut and determined a mean competition ratio for a typical larval gut (competitor1:competitor2). Following gut dissections, EM exchanges were repeated every two hours to ensure that bacterial loads in the EM remained below 10^3 CFU/mL (monitored by plating a single 100 μL sample just prior to the beginning of each set of EM exchanges). The larvae that remained in the competition petri dishes after the first set gut dissections were completed, were incubated at 28 ˚C between sterile EM exchanges. 8.5-10.5 hours after initial gut dissections took place, 10 additional larval guts were dissected, and their contents dilution plated to establish a final mean competition ratio for a typical larval gut (competitor1:competitor2). Since the CFU/mL in the EM was maintained at a low level between bouts of dissection, any increase in the mean CFU/gut

I observed between dissection bouts should have been affected primarily by the growth of *S. oneidensis* populations *in vivo*. Using the initial mean CFU/gut and the final mean CFU/gut that I observed for each competing strain, I then calculated a per capita growth rate for each strain per competition ($r = \ln\left(\frac{CFU/gut_{final}}{CFU/gut_{initial}}\right)$). I then used my per capita growth rate to calculate an immigration index and relative *in vivo* fitness metric: (immigration index: $\frac{gut\ ratio_{competitor1:competitor2}}{inoculation\ ratio_{competitor1:competitor2}} * e^{(r_{competitor2} - r_{competitor1})*t}$; relative *in vivo* fitness: $\frac{r_{competitor1}}{r_{competitor2}}$). In this way, my *in vivo* relative fitness compares the relative per capita growth rates I observed *in vivo* for each competitor, and I was able to use these per capita growth rates to control for any growth that may have occurred *in vivo* during my immigration assay.

Motility assays:

5 mL overnight TSB cultures of strains of interest were diluted 1:100 in TSB and allowed to grow out to late log stage (4-5 hours). Strains were then prepped for inoculation as described in my whole experimental system competitions above (combine differentially fluorescently-tagged competitors from subcultures, pellet, resuspend in sterile EM, and dilute). 7.5 μL of each prepared competition mixture were then used to inoculate BF larval flasks containing ~15 mL of EM and ~15 larvae at 4 dpf. Inoculated larval flasks were incubated at 28 ˚C for 13-17 hours, at which point bacteria in each flask were imaged on an inverted microscope (Nikon Eclipse Ti-e) by focusing on the bottom interior surface of the flask. Ten, 30 second movies were then captured separately for each tagged population in a competition flask at a rate of > 15 frames per second. To

calculate motility characteristics such as the speed of motile cells and the proportion of a population that was motile, I relied on a particle tracking algorithm (Parthasarathy, 2012; source code: http://pages.uoregon.edu/raghu/particle_tracking.html). Only cells deemed to be moving faster than 2 um/second were considered motile.

Results:

Shewanella oneidensis is not host associated:

Gut-associated bacteria are routinely isolated from the guts of their animal hosts. Although MR-1 has never been observed within a host gut, the *Shewanella* genus has been commonly found in larval zebrafish gut (Stephens et al., 2016). Indeed MR-1 shares a recent common ancestry with a *Shewanella* species that has recently been isolated from the gut (Figure 2, A & B; *Shewanella* ZOR0012, Shew-Z12 from this point forth). Interestingly, a whole genome comparison between MR-1 and Shew-Z12 reveals that these two genomes share an average nucleotide identity (ANI) of approximately 89% (see methods; Yoon et al., 2015). The high degree of overlap between these two species was further reflected when I compared the protein sequence alignments of Shew-Z12 or the more distantly related *Shewanella woodyi* against my MR-1 reference genome (Figure 2B). All three of these species are classified within the same genus, yet relative to *S. woodyi*, MR-1 appears to display much higher levels of amino acid sequence identity with Shew-Z12 on a per gene basis, implying greater amounts of functional conservation between MR-1 and Shew-Z12.

Given the close relation between MR-1 and Shew-Z12, I thought it prudent to ensure that MR-1 would not behave like a host-associated bacterium. Thus, I compared

its fitness within the zebrafish gut to that of Shew-Z12 via two separate metrics: 1) colonization density in the gut after 72 hours exposure and 2) competitive ability to colonize the gut. Despite being able to colonize BF larvae under monoassociation conditions, MR-1 did so at a lower level than Shew-Z12 under these same conditions (Figure 3A). Additionally, compared to how well an MR-1 strain with a neutral fluorescent tag (MR-1tag) competed against a wild type (wt) MR-1 strain (MR-1wt), which resulted in a tie, the MR-1tag strain was severely outcompeted by the Shew-Z12 strain in its ability to colonize the BF larval guts (Figure 3B, see methods for a description of how these competitions were conducted). Together, these lines of evidence suggest that MR-1 is not well-adapted for life in association with the larval zebrafish gut. Thus, I speculated that, depending on how much these two organisms have diverged from their shared common ancestor, adapting MR-1 to the host environment from which Shew-Z12 was isolated might elucidate the adaptive steps Shew-Z12 took to become host associated.

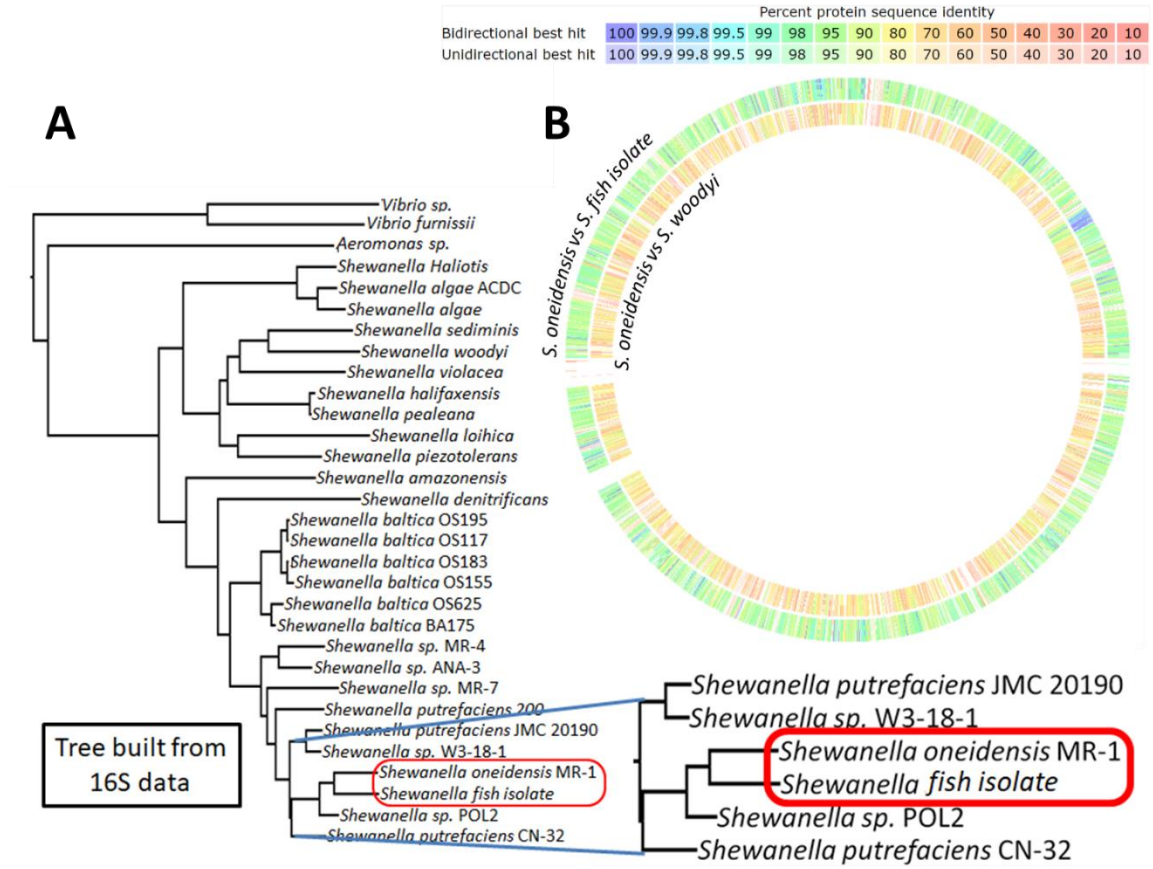

Figure 2: Relatedness of MR-1 to other *Shewanella* species. A) A phylogenetic tree showing the relationship between a number of different *Shewanella* species. The tree was compiled using the 16S gene of each species represented. B) Whole genome alignment of Shew-Z12 (outer ring) or *Shewanella woodyi* (inner ring) amino acid sequences against the MR-1 reference genome. Color shading around each ring indicates sequence identity (key at the top of the figure). Alignments were conducted using the sequenced-based comparison tool of The SEED Viewer v. 2.0 (Overbeek et al. 2013).

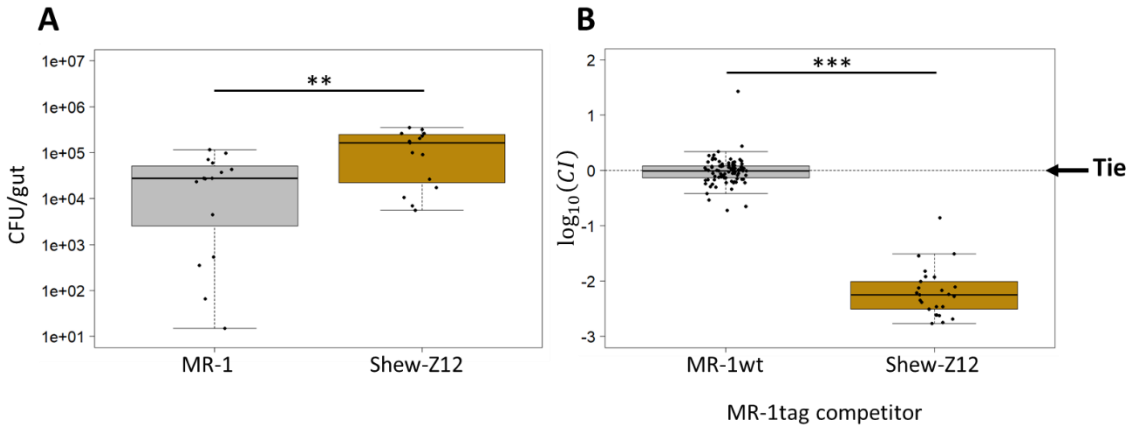

Figure 3: Fitness comparison between MR-1 and Shew-Z12. A) Colonization density achieved in larval guts after 72 hours of colonization under monoassociation conditions for indicated strains. Dissected guts were plated on TSA and colony-forming units (CFUs) were counted. Each point represents a single dissected gut. B) Competitive ability of unpassaged fluorescently-tagged *MR-1* strain competing against an untagged version of itself (MR-1wt), or an untagged Shew-Z12. Each point represents the competitive index (CI) measured for a single larval gut (see Methods for details on how Cis were calculated).

Serial passage should uncover how MR-1 adapts to the gut:

To understand how *Shewanella oneidensis* would adapt to a vertebrate host gut I serially passaged six replicate populations through the digestive tracts of GF larval zebrafish (Figure 1A). At the start of the experiment, each population was composed of two isogenic fluorescently-tagged MR-1 isolates (dTomato, MR-1dT; gfp, MR-1gfp) so that passaged populations could be distinguished from the unpassaged ancestor by their colony color on TSA agar plates, and adaptive events could be inferred from changes in each tag's frequency within an evolving population. After 20 passages I tested the fitness of at least one isolate per evolved population by competing each isolate against their unpassaged ancestor. Each competition involved inoculating larval flasks containing 10-15 larvae with competition mixtures composed of roughly equal proportions of MR-1wt and each evolved isolate at a density of ~10^3 CFU/mL. Evolved:ancestor inoculation

ratios were determined by plating triplicate samples of the embryonic medium (EM) from each competition flask and quantifying the number of fluorescently tagged (evolved) and untagged (ancestor) CFUs. To mimic a standard passage from my serial passage protocol, I then incubated larvae with competition mixtures for 72 hours at which point I dissected and plated 10 larval guts. Finally, I calculated competitive indices (CIs) by dividing the evolved:ancestor CFU ratio of each plated gut sample by the mean inoculation ratio from my triplicate inoculation assessment. Five of the six isolates tested showed significant improvements in their ability to outcompete the ancestral strain (Figure 4A), and these improvements were not based on adaptation to the general lab environment as competitions between whole replicate evolved populations and MR-1wt in rich media produced CIs around zero after log transformation (Figure 4B).

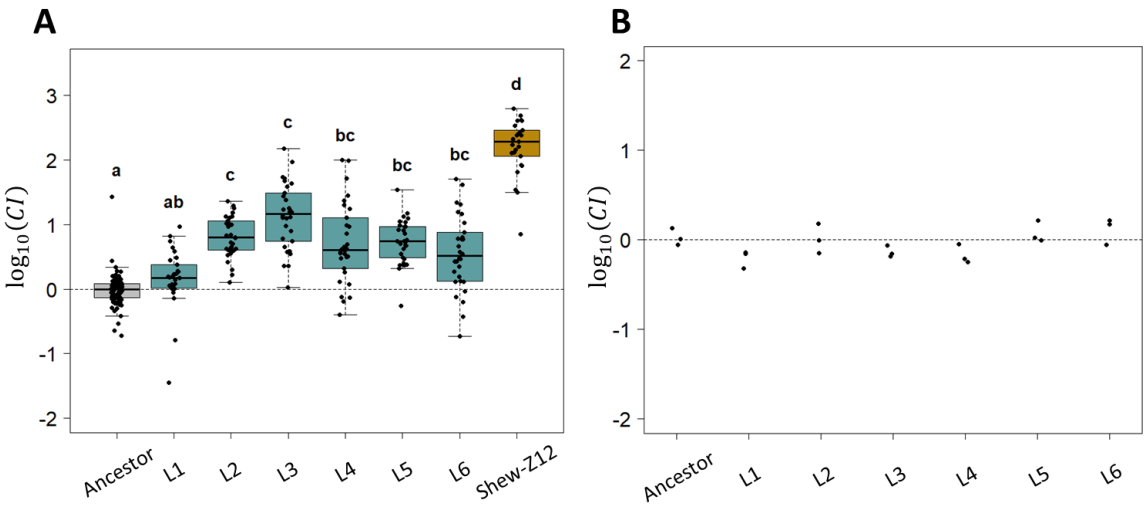

Figure 4: Competitive fitness of evolved *S. oneidensis* isolates. A) Competitive ability of MR-1 isolates from each replicate evolved population against the ancestral MR-1 reference strain. Each point represents the CI measured for a single larval gut. An ancestral competition against itself (left-most box) is shown as a control to represent the absence of a competitive advantage. Statistical groupings are indicated by letters above each box for a significance threshold of $p<0.05$. Letters in common between groups indicate the absence of a significant difference in each group's mean. B) Competitive ability of each replicate evolved MR-1 population competing against the ancestral MR-1 reference strain in tryptic soy broth (TSB). Each point represents the CI measured for a single biological replicate (see Methods for details on how CIs were calculated). Each TSB competition was performed in 10 mL TSB in a 20 mL glass test tube.

Several adaptive candidate mutations were found using comparative genomics:

Next, to determine the genetic underpinnings that might explain the adaptive advantages of each evolved isolate tested, I sequenced the genomes of four isolates per evolved population (24 isolates total). The reads of each sequenced isolate were then aligned against an ancestral reference sequence to identify regions of the evolved genomes that had accumulated mutations. I relied on breseq 0.31.0 in consensus mode to conduct these comparisons. The resulting list featured a large number of potential adaptive candidate loci that had undergone selection (Table 1) which made it difficult to distinguish adaptive mutations from non-adaptive mutations that co-occurred in in the

same genome. For example, mutations with neutral or even slightly deleterious impacts on fitness could have been selected to increase in frequency simply through their linkage with another, more beneficial mutation. Since it would be unlikely for mutations to accumulate in similar genomic regions across separate evolved isolates solely due to chance events and drift, I hypothesized that genomic regions that contained high densities of mutations had likely been selected to increase in frequency due to their elevated adaptive value. Thus, I narrowed my scope of adaptive candidates by looking for similar types of mutations across all sequenced isolates. In 20 out of 24 isolates sequenced, I observed mutations in a mannose sensitive hemagglutinin (msh) pilus operon (mshOP; Table 1, yellow shaded).

Table 1: Candidate adaptive mutations. Mutations are listed for four isolates per evolved population. Each isolate is color-coded by the fluorescent tag it contained (red = dTomato, green = gfp). Except for mshOP annotations (annotated with RAST v. 2.0), gene annotations are based on Prokka v1.12 annotations. Color-coding for gene descriptions is as follows: green: synonymous mutation, blue: missense mutation, red: nonsense mutation, black: all other mutations. Yellow-shaded cells indicate mshOP mutations.

Line 1				Line 2				Line 3				Line 4				Line 5				Line 6							
A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	Gene	Mutation	Description	Annotation
										X	X													*mshL* ←	C→A	E75*	MSHA biogenesis protein MshL
		X	X			X	X	X	X					X	X									*mshL* ←	T→G	T300P	MSHA biogenesis protein MshL
X	X											X	X											*mshG* ←	G→T	A289E	MSHA biogenesis protein MshG
																				X	X	X	X	*mshE* ←	G→A	Q196*	MSHA biogenesis protein MshE
																X								*mshE* ←	C→A	E83*	MSHA biogenesis protein MshE
				X																				*mshO* ←	Δ1 bp	541/849 nt	MSHA biogenesis protein MshO
		X																						*00112* →	G→A	E27K	hypothetical protein
												X												*00235* ←	G→T	G137G	Sulfite exporter TauE/SafE
								X												X				*00301* →	A→G	K37R	hypothetical protein
								X												X				*00301* →	G→T	T41T	hypothetical protein
								X												X				*00301* →	G→A	E50E	hypothetical protein
								X												X				*00301* →	T→G	L52L	hypothetical protein
								X												X				*00301* →	G→A	E65E	hypothetical protein
								X												X				*00301* → / ← *00302*	T→A	intergenic (+5/+81)	hypothetical protein/Suppressor of fused protein (SUFU)
																X	X	X						*rcsC_1* →	G→A	A304T	Sensor histidine kinase RcsC
									X								X	X		X	X	X		*00448* →	C→A	Q182K	Transposase DDE domain protein
									X								X	X		X	X	X		*00448* →	C→T	I215I	Transposase DDE domain protein
									X								X	X		X	X	X		*00448* →	T→C	T285T	Transposase DDE domain protein
									X								X	X		X	X	X		*00448* →	C→T	N293N	Transposase DDE domain protein
																	X	X		X	X	X		*00448* →	A→C	V154V	Transposase DDE domain protein
																	X	X		X	X	X		*00448* →	C→A	V171V	Transposase DDE domain protein
		X																						*00582* ← / ← *proS*	C→A	intergenic (-167/+109)	Patatin-like phospholipase/Proline--tRNA ligase
							X																	*00712* ←	T→C	N23D	hypothetical protein
								X				X								X				*00712* ←	2 bp→TT	85-86/228 nt	hypothetical protein
								X				X								X				*00712* ←	T→C	N23D	hypothetical protein
								X				X								X				*00712* ←	C→T	L20L	hypothetical protein
								X				X								X				*00712* ←	A→G	G18G	hypothetical protein
				X																				*dsbA_2* → / → *00890*	G→A	intergenic (+778/-142)	Thiol:disulfide interchange protein DsbA precursor/hypothetical protein
						X			X									X	X		X	X	X	*00937* ←	2 bp→GC	59-60/333 nt	hypothetical protein
						X			X		X							X	X		X	X	X	*00937* ←	A→T	V13E	hypothetical protein
						X			X		X							X	X		X	X	X	*00937* ←	A→G	I9I	hypothetical protein
						X			X		X							X	X		X	X	X	*00937* ←	C→T	R7H	hypothetical protein
						X			X		X							X	X		X	X	X	*00937* ←	C→T	M6I	hypothetical protein
						X			X									X	X		X	X	X	*00937* ← / ← *00938*	C→T	intergenic (-11/+137)	hypothetical protein/plasmid segregation centromere-binding protein ParG

Table 1 continued

Line 1				Line 2				Line 3				Line 4				Line 5				Line 6				Gene	Mutation	Description	Annotation
A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D				
									X									X	X		X	X	X	*00937 ← / ← 00939*	2 bp→TG	intergenic (-38/+109)	hypothetical protein/plasmid segregation centromere-binding protein ParG
X	X											X	X											*glpC →*	T→G	F257C	oxidoreductase, FAD-binding, putative
																	X							*01176 →*	T→G	W460G	lipoprotein, putative
							X																	*01736 ←*	A→G	F149S	Zinc carboxypeptidase
																					X			*folD → / ← cysS*	$(TAGGTTC)_{14\rightarrow13}$	intergenic (+89/+81)	Bifunctional protein FolD protein/Cysteine--tRNA ligase
																	X	X	X					*dosC →*	$(A)_{6\rightarrow5}$	375/1557 nt	Diguanylate cyclase DosC
				X	X																			*gmr_6 →*	G→A	A174T	Cyclic di-GMP phosphodiesterase Gmr
								X	X															*smc_2 ←*	A→G	M446T	Chromosome partition protein Smc
																	X	X	X					*02921 →*	$(A)_{8\rightarrow7}$	1418/2091 nt	hypothetical protein
X																								*secY ←*	G→C	P355A	protein translocase subunit secY/sec61 alpha
									X									X	X		X	X	X	*03772 ←*	2 bp→GG	372-373/444 nt	hypothetical protein
									X									X	X		X	X	X	*03772 ←*	A→G	D123D	hypothetical protein
									X									X	X		X	X	X	*03772 ←*	C→T	S122N	hypothetical protein
									X									X	X		X	X	X	*03772 ←*	T→G	M121L	hypothetical protein
									X									X	X		X	X	X	*03772 ←*	A→G	S117P	hypothetical protein
									X									X	X		X	X	X	*03772 ←*	G→A	L115L	hypothetical protein
									X									X	X		X	X	X	*03772 ←*	G→A	S112S	hypothetical protein
									X									X	X		X	X	X	*03772 ←*	T→A	L110L	hypothetical protein
									X									X	X		X	X	X	*03772 ←*	A→C	T109T	hypothetical protein
									X									X	X		X	X	X	*03772 ←*	G→A	V106V	hypothetical protein
									X									X	X		X	X	X	*03772 ←*	A→T	G99G	hypothetical protein
								X																*04242 →*	2 bp→TT	50-51/576 nt	Integrase core domain protein
								X																*04242 →*	C→A	P25T	Integrase core domain protein
								X																*04242 →*	T→C	P25P	Integrase core domain protein
								X																*04242 →*	A→C	A27A	Integrase core domain protein
								X																*04242 →*	T→A	F36I	Integrase core domain protein

While there were multiple types of mutations that were found in separate evolved populations, no other class of mutation was as widespread as those found in the mshOP (all six replicate populations featured isolates with mshOP mutations). Therefore, to assess whether these mutations provided evidence that MR-1 was perusing a similar adaptive trajectory to one potentially taken by Shew-Z12, I compared the mshOP of my reference MR-1 strain to that of Shew-Z12, to look for any indication that MR-1 might be experiencing genetic convergence toward Shew-Z12. Despite high levels of sequence similarity overall (ANI: 91.64%), an alignment of the Shew-Z12 mshOP against that of MR-1 revealed several regions with elevated patterns of divergence (Figure 5). In

particular, Shew-Z12's mshQ gene appeared to be noticeably larger than MR-1's mshQ gene. Additionally, there also appeared to be a region of high divergence toward the N-terminus of the MshA protein between these two species, and MshA is the protein that forms the major pilin subunit of the msh pilus (Jones et al., 2015). However, I did not find mutations in the MR-1 mshA gene in any of my evolved isolates. Nonetheless, the fact that there were substantial regions of divergence between the mshOP amino acid sequences of MR-1 and Shew-Z12 left open the possibility that the adaptive changes I observed might amount to evolutionary convergence at the functional level.

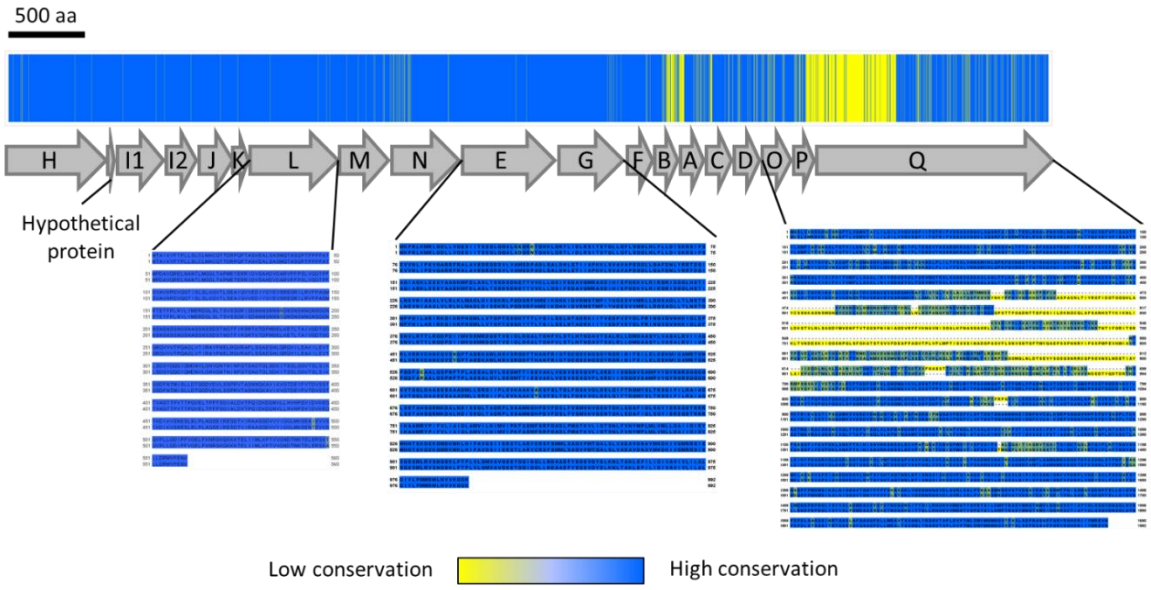

Figure 5: **MshOP amino acid conservation between MR-1 and Shew-Z12.** The bar at the top shows the mshOP amino acid conservation between MR-1 and Shew-Z12. A schematic depicting the organization of the mshOP is shown below. Per site, blue indicates the same amino acid is present in both species, while yellow indicates the presence of different amino acids. Aligned sequences are shown for genes in which I observed mutations in my evolved isolates. In each row of the displayed sequences, MR-1 is featured on top and Shew-Z12 is featured on the bottom. MshQ did not accumulate mutations in my experiment, but it is shown to illustrate that much of the divergence in this protein resulted from a difference in length between MR-1 and Shew-Z12. The color coding in each sequence alignment indicates the degree of conservation. Amino acids with similar biochemistry are bluer, while those with divergent biochemistry are yellower. A scale bar above the figure indicates the length of 500 amino acids.

Most of the other classes of mutations that were found in multiple evolved isolates occurred in genes that were annotated as hypothetical proteins or transposases (Table 1). Given these annotations, it was difficult to hypothesize potential mechanisms for how alterations in these genes might enhance competitive fitness. For instance, the term "hypothetical protein" provides no functional information, and mutations in transposases would presumably retarget those transposases to recognize different sequences. In turn, these transposases could then modify other genes in the genome, and thus their impact on fitness would be indirect. Alternatively, mutating transposases, which typically bind DNA, could prevent them from binding their targets, allowing access by transcription factors and other DNA binding proteins, that could affect gene expression. Either way, while I could not rule out their potential impact on fitness, considering mutations in these genes provided no obvious link to fitness, and thus these mutations were not examined further.

Interestingly, the mshOP has been implicated in a number of other host microbe systems. Specifically, within the *Vibrio* genus, msh pili, which are hair-like structures that extend from bacterial cell surfaces (Figure 6, A and B), have been suggested to modulate various *Vibrio* species' ability to colonize and interact with the corporal environments of both vertebrate and invertebrate hosts (Hsiao et al., 2006; List et al., 2018; Ariyakumar & Nishiguchi, 2009; O'Boyle et al., 2013). In an infant mouse model, *V. cholerae* msh pili were implicated in their ability to alter colonization dynamics and were suggested to interact with a component of the mouse immune system, IgA (Hsiao et al., 2006). Additionally, a *Pseudomonas aeruginosa* strain that expresses msh pili on its surface (PA-MSHA) has also been shown to interact with mammalian immune systems

(Hou et al., 2012). Given is body of work, the fact that mshOP mutations commonly arose and were maintained in each of my replicate evolved populations, and the fact I observed ample amounts of divergence between the mshOPs of MR-1 and Shew-Z12, I decided to focus on how the most frequently observed msh operon mutation (a threonine to proline missense mutation at amino acid 300/560; MshL-T300P) impacted fitness.

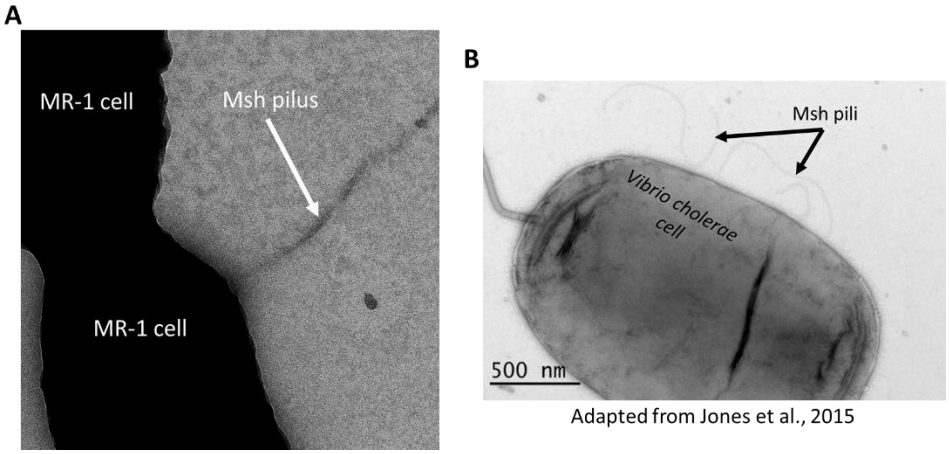

Figure 6: **Transmission electron microscopy (TEM) images of msh pili.** A) I generated a TEM image of a dividing MR-1 cell. The cell on the bottom half of the image is expressing an msh pilus. The details regarding how I generated this image can be found in the methods section. B) A TEM image of a *Vibrio cholerae* cell expressing msh pili is shown for comparison. This image was adapted from Jones et al., 2015.

MshL-T300P mimics MshL deletion mutant fitness:

To confirm whether the MshL-T300P mutation could improve fitness, I used allelic exchange to recreate this mutation in the ancestral genomic background (method described in Wiles et al., 2018). I then assayed the fitness of this mutant using a competition assay. The MshL-T300P mutant significantly outcompeted the wt ancestor implying that this mutation is responsible for the increased fitness seen in the evolved isolates in which it was observed (Figure 7). Given the different character of threonine

and proline side chains, and how early in the mshL gene this missense mutation occurred, I hypothesized that the MshL-T300P mutation was a loss of function mutation. To confirm this, I created a mshL whole gene knockout (ΔmshL) and competed it against the MR-1 ancestor. The ΔmshL mutant also showed improved fitness that was comparable to that of the MshL-T300P mutant, implying that the MshL-T300P mutation was indeed a loss of function mutation (Figure 7). Finally, to assess whether the MshL-T300P mutation was the primary driver of increased fitness in evolved isolates containing this mutation, I competed the L3 isolate against the MshL-T300P mutant. I hypothesized that if the MshL-T300P mutant imparted an increase in fitness that was of comparable adaptive value to the evolved L3 isolate which contained this mutation, this competition should result in a dramatic reduction in L3's competitive index. This is indeed what I observed (Figure 7, rightmost column), and given that the L3 isolate had the highest median fitness of all tested isolates containing msh operon mutations, this result implies that the MshL-T300P mutation imparts a large adaptive effect.

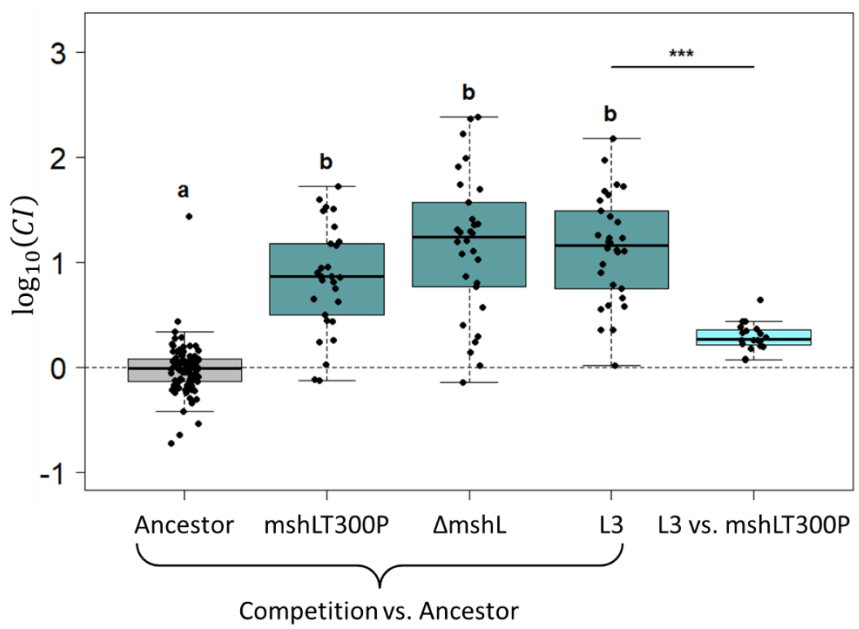

Figure 7: The competitive ability of the MshL-T300P mutant, ΔmshL mutant, and L3 isolate against MR-1wt. Each point represents the competitive index (CI) measured for a single larval gut. Ancestral competition against itself is shown as a control to represent the absence of a competitive advantage. Statistical groupings are indicated by letters above each box for a significance threshold of $p<0.05$. Letters in common between groups indicate the absence of a significant difference in each group's mean. Statistics for the L3 vs MshL-T300P competition (right-most box) were conducted separately since this competition had unique competitors ($p<0.001$: ***).

Losing mshL functionality reduces biofilm formation on surfaces:

The msh pilus has been shown to be important for surface attachment and biofilm formation in MR-1, *V. cholerae*, and in *Aeromonas hydrophila* (Saville et al., 2010; Bouhenni et al., 2010; Jones et al., 2015; Qin et al., 2014). To assess whether MshL-T300P mutants could form normal biofilms, I conducted a static polystyrene plate-based biofilm assay in larvae conditioned medium (LCM). This medium is created by allowing zebrafish larvae to condition the embryonic medium (EM) used to derive them BF for 96 hours, and it should provide a similar nutrient profile to that experienced by MR-1 during the competitive fitness assays described above. I found that the L3 isolate as well as both the MshL-T300P and ΔmshL mutants formed reduced biofilms compared to the ancestral

MR-1 strain (Figure 8). These results imply that the MshL-T300P corrupts the ability to properly adhere to surfaces, and they provide further confirmation that MshL-T300P is a loss of function mutation. Since previous studies had shown that the expression of msh pilus was required for normal biofilm formation, these results are consistent with the interpretation that the MshL-T300P mutation reduces the expression of the msh pilus on the surface of MR-1 cells. If true, it would be improbable that the putative loss-of-function MshL-T300P mutation featured in L3 represents an example of functional convergence between MR-1 and Shew-Z12, since Shew-Z12 appears to have a fully intact mshOP that shares high sequence identity with MR-1.

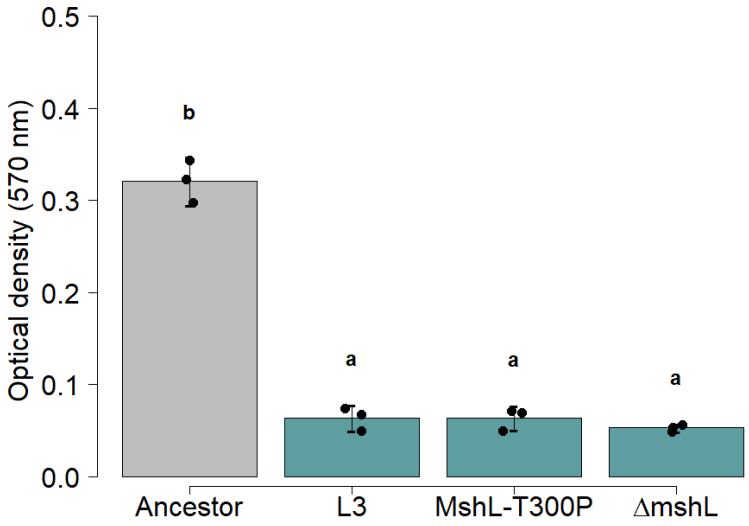

Figure 8: Crystal violet biofilm assay comparing L3, MshL-T300P, and ΔmshL to ancestor. Optical density (570 nm) corresponds to crystal violet intensity. Higher optical density readings indicate more robust biofilms. Statistical groupings are indicated by letters above each box for a significance threshold of $p<0.05$. Letters in common between groups indicate the absence of a significant difference in each group's mean.

Pili have also been shown to be important for twitching motility whereby bacteria extend pili, attach to a surface, and then retract the pili resulting in a slow crawling-type of locomotion across a substrate. Based on my biofilm results which suggested that msh-pilus mediated biofilm formation was hampered in mshL mutants, I hypothesized that

MshL-T300P mutants would lack twitching motility. I found no evidence of twitching motility in either the wt MR-1 ancestor or the mshOP mutants (data not shown), and this is consistent with other studies that have found MR-1 does not twitch (Saville et al., 2010).

Losing mshL functionality has no effect on carrying capacity in larval guts:

After gaining some mechanistic insight into how the MshL-T300P mutation impacts MR-1 physiology, I sought to understand how this mutation improved fitness. Since the Shew-Z12 zebrafish isolate demonstrated the ability to colonize BF larval guts at greater densities than MR-1, I examined whether the adaptive MshL-T300P mutation could improve the total abundance reached in the gut when the ancestor was competing against itself, versus when the ancestor was competing against either the L3 isolate or the MshL-T300P mutant. If the MshL-T300P mutation had a discernible impact MR-1's carrying capacity in the gut, I expected to observe greater cell densities in competitions where MshL-T300P-containing strains were present. My data show that when the L3 and MshL-T300P strains were used as competitors, there was no significant change in carrying capacity (Figure 9). This suggests that the evolved L3 isolate's improved fitness was not based on an ability to colonize larval guts at higher abundances.

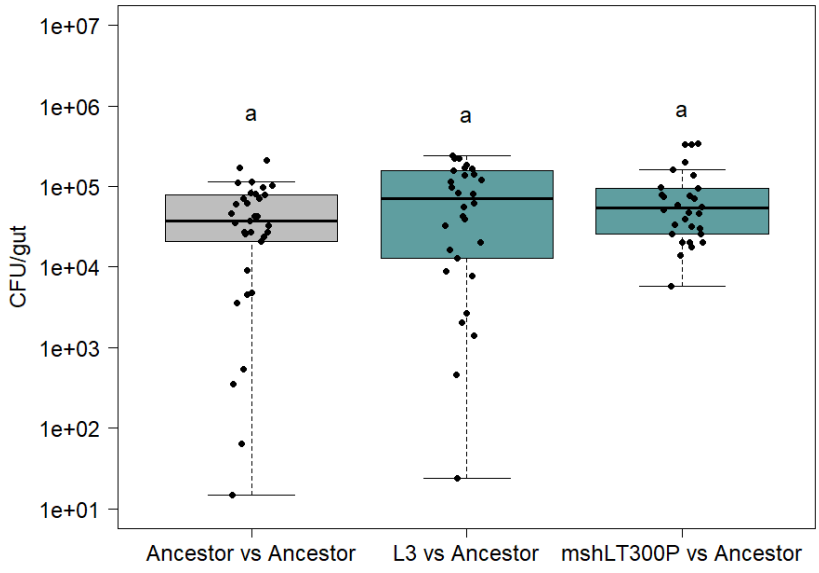

Figure 9: Colonization density achieved in larval guts after 72 hours of colonization under competitive conditions for indicated strains. Dissected guts were plated on TSA and CFUs were counted. Each point represents a single dissected gut. Statistical groupings are indicated by letters above each box for a significance threshold of $p<0.05$. Letters in common between groups indicate the absence of a significant difference in each group's mean.

L3 and T300P have similar in vivo growth rates to the ancestor:

Given that I did not observe differences in carrying capacity that could explain the fitness advantage I observed for the L3 and MshL-T300P strains in my system, I hypothesized that the MshL-T300P mutation must impart its advantage through other components of fitness. In my experimental system, in order for bacterial cells to be propagated from one passage to the next, populations must grow to a density where they can readily colonize larvae, bypass host filters to migrate from the aqueous environment into the larval digestive tract, grow within the digestive tract, and resist ejection from the digestive tract such that they are present in the gut at the time of dissection. Isolates that have increased their ability to perform any of these tasks relative to the ancestral strain, should register as having increased fitness. Fortunately, part of the power of the zebrafish

system featured here, is that it allows for the isolation of several of these components of fitness, enabling us to assess their individual contribution to the overall fitness I observe. In this way, evolution can act as a sensor, elucidating which aspects of the experimental environment provide the most fodder for adaptation.

I first suspected that my evolved populations might have improved in their ability to compete against the ancestral MR-1 strain *in vivo* once they had colonized larval guts. To test this hypothesis, I performed an *in vivo* growth experiment in which I dissected the guts of groups of 10 larval zebrafish colonized with competition mixtures of L3 or MshL-T300P versus the ancestral MR-1 strain at two different time points. 10 larval guts were dissected at an initial time point, when populations were approximately two orders of magnitude below the carrying capacity of the larval gut, to establish a mean founding population size for each competitor in a typical larval gut. 10 additional larval guts were then dissected ~10 hours later, after some growth in the gut had taken place, to establish a mean final population size per typical gut. While my gut sampling scheme is destructive, preventing us from us from calculating growth rates of MR-1 populations within a single larval gut, based on the average initial and final populations I observe, I estimate that bacterial populations in an average larval gut undergo roughly 3.3 doublings during this 10-hour window. This provides ample opportunity for competing strains with different levels of fitness to differentiate themselves. Additionally, based on previous work conducted with other evolved isolates, I determined that larval guts were not colonized at cell densities below 10^4 CFU/mL (Figure 10, A and B). Therefore, I regularly exchanged the EM in my *in vivo* growth experiment to ensure that all cells detected in the gut at the end of my experiment were the result of growth alone. At the conclusion of this

experiment, I was able to calculate a relative fitness metric based on the per capita growth rates of each competitor ($W_{relative} = \frac{\ln\left(\frac{competitor1_{final\ CFU/gut}}{competitor1_{initial\ CFU/gut}}\right)}{\ln\left(\frac{competitor2_{final\ CFU/gut}}{competitor2_{initial\ CFU/gut}}\right)}$; Lenski et al., 1991). I found that neither the L3 isolate nor the MshL-T300P mutant outcompeted the ancestor *in vivo*, and surprisingly, it appears that the MshL-T300P isolate may even under perform the ancestor (Figure 10C).

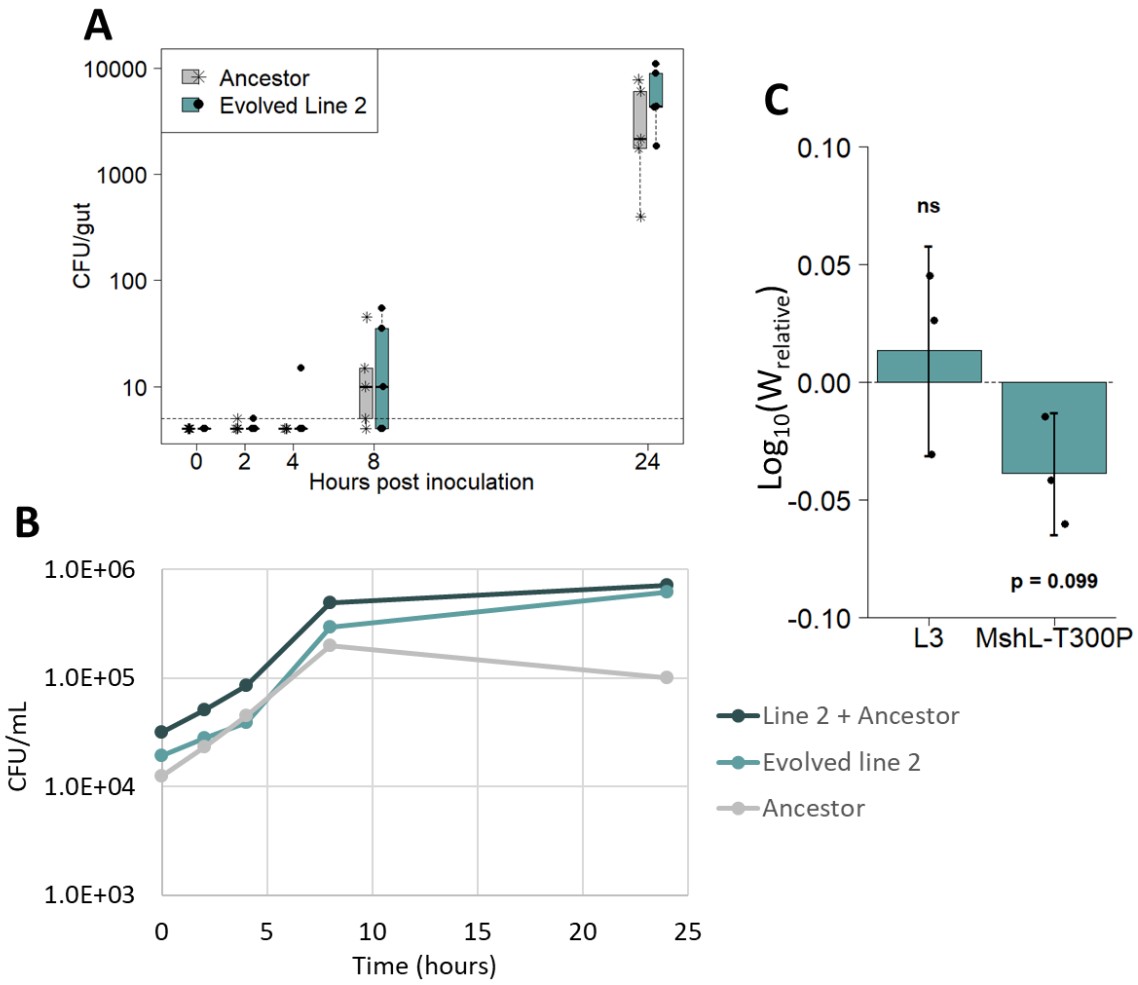

Figure 10: Competitive fitness of L3 and MshL-T300P mutant *in vivo.* A) Colonization density over time. A larval BF larval flask was inoculated with competition mixture containing evolved population 2 (blue boxes) and MR-1wt (gray boxes). Five larval guts were then dissected and plated at each of the time points indicated. CFU/gut was quantified for each strain in each dissected gut, and each point shown represents the colonization density of a single gut. The dotted line indicates the limit of detection. B) Same experiment as in A) except that cell densities (CFU/mL) were measured in the EM outside the larvae over the same period. C) Competitive *in vivo* growth rates were calculated as described in the text and then used to determine a relative *in vivo* fitness metric ($W_{relative}$). Each point represents the log-transformed mean $W_{relative}$ for a single competition. The $\log_{10}(W_{relative})$ for each group of points was t-tested against the value zero, my expected value if there were no competitive advantage.

L3 and T300P outcompete the ancestor in LCM:

Because my adaptive strains exhibited no ability to outcompete he ancestral MR-1 strain *in vivo*, I next sought to interrogate how much of the L3 isolate's fitness was based solely on its adaptation to the external environment (i.e. the medium that has been conditioned by larval zebrafish cohorts). Thus, I conducted competition assays in LCM, devoid of larval hosts, that was generated as described above in my biofilm assays. In these LCM competitions, I established the mutant:ancestor ratio of an inoculating competition mixture in triplicate by counting the CFUs from each competing strain on agar plates. I then incubated the competition for 72 hours, and again assessed the mutant:ancestor ratio by CFUs on plates at the end of the experiment. CIs were calculated by dividing the final ratio by the initial ratio as before. In these relatively simple assays, acquiring samples necessitates withdrawing fluid from the water column. Thus, the only way one strain could outcompete another is by increasing its frequency specifically in the planktonic portion of the environment (where the larvae would normally reside), rather than near the surface of the flasks in which these competitions are performed. I found that both the MshL-T300P mutant and the L3 isolate were able to outcompete the ancestral MR-1 strain, implying evolved and mutant these strains had greater representation in the water column (Figure 11A).

I additionally monitored the cell densities over time for LCM competitions between the MshL-T300P mutant and the MR-1 ancestor in triplicate and found that both cell populations reached their peak densities in the water column at around 17 hours post inoculation (MshL-T300P: 2.2e5$\pm$5.9e4, MR-1wt: 5.5e4$\pm$1.3e3; mean $\pm$ SD; Figure 11B), and that the MshL-T300P:ancestor ratio appeared to reached its highest point

around 35 hours post inoculation ($\log_{10}$(MshL-T300P/Ancestor) = 1.14±0.46; mean ± SD; Figure 11C). Interestingly, when the ancestral MR-1 strain was cultured in LCM by itself, it reached a peak density that was higher than the combined total reached in the MshL-T300P versus ancestor experiment, at roughly the same time point (Figure 11D). Although these growth curves were measured in triplicate on different days (meaning that the nutrient content in these experiments could have been slightly different), this suggests that the ancestor can grow at a similar rate to that of the MshL-T300P mutant in the absence of this more fit competitor. Because the ancestral population had attenuated growth under competitive conditions, together these results are consistent with the interpretation that the MshL-T300P mutant is better able to assimilate nutrients within the LCM compared to the ancestor under competitive conditions.

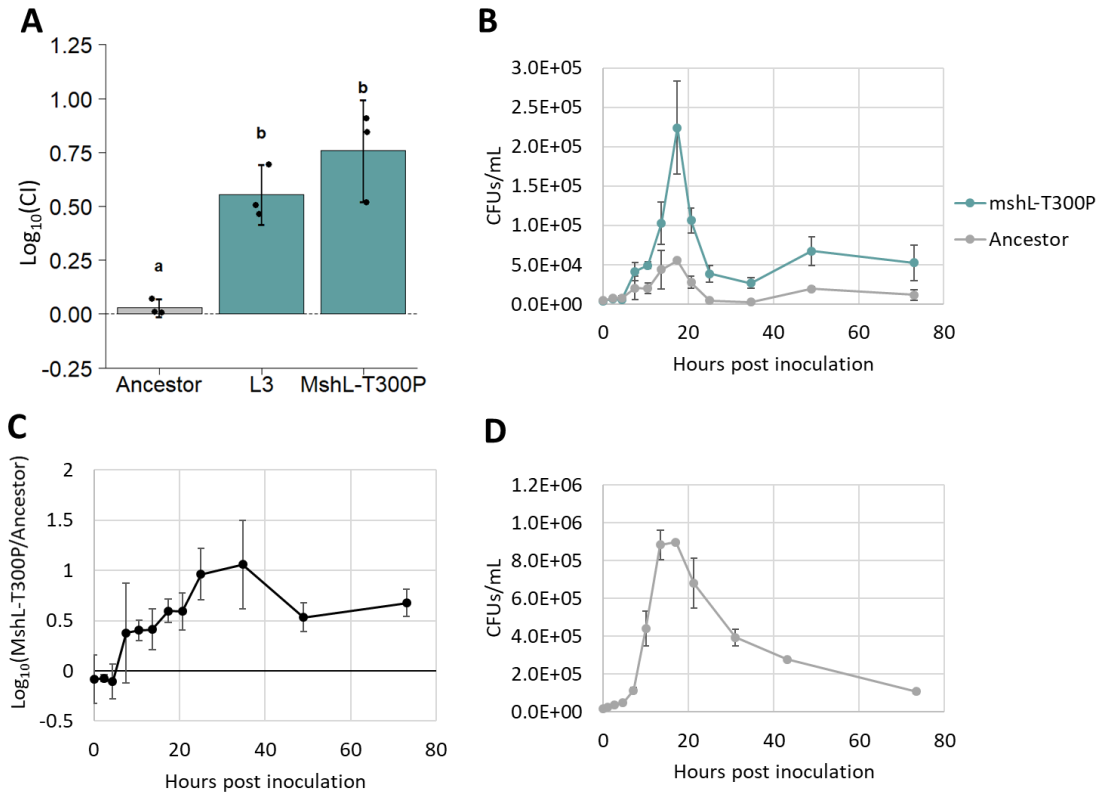

Figure 11: Competitive fitness of L3 and MshL-T300P mutant in larvae-conditioned media. A) Competitive ability of L3 isolate and the MshL-T300P mutant against MR-1wt in LCM. Each point represents the competitive index (CI) measured for a single LCM competition flask. Ancestral competition against itself is shown as a control to represent the absence of a competitive advantage. Statistical groupings are indicated by letters above each box for a significance threshold of $p<0.05$. Letters in common between groups indicate the absence of a significant difference in each group's mean. B) Population dynamics over time in LCM during a competition between MshL-T300P and MR-1wt. C) MshL-T300P:MR-1wt ratio over the course of the experiment shown in B). D) Population dynamics over time in LCM for MR-1wt growing by itself. For B), C), and D), error bars indicate the standard error of the mean for three biological replicates.

L3 and mshLT300P increase migration into the gut:

Although the evolved and mutant strains are likely able to outcompete the ancestor based in part on their greater representation in the planktonic phase of my experimental setup, I wondered whether these strains could also bypass host filters to immigrate into the gut more efficiently on a per capita basis. To assess this, I first

cultured competitor strains separately in LCM for 12-15 hours, and then created competition mixtures by combining the fluid from either the L3 isolate or MshL-T300P cultures with the fluid from an ancestral culture. The 12-15-hour culturing period was chosen to coincide with MR-1 cell densities that are near their peak abundance (Figure 11, B and D), at which point MR-1 can readily colonize BF larvae (Figure 10, A and B). I then exposed BF larvae to this competition mixture and immediately sampled the media to quantify the initial ratio of mutant or evolved cells to ancestral cells. 40-60 minutes after larval exposure to MR-1 competition mixtures, 10 larval guts were dissected, homogenized, and plated to assess the ratio of evolved or mutant cells to ancestral cells that had colonized each gut. I then calculated an immigration index for each larva by dividing the colonizing ratio by the initial ratio and accounting for any *in vivo* growth differences between the two competing strains that may have biased my results (see methods for details). For both the L3 and MshL-T300P competitions against the ancestor, the mutant and evolved strains out competed the ancestor in terms of their per capita colonization of the larval gut (Figure 12). This result suggests that the MshL-T300P mutation can increase the rate at which MR-1 cells migrate from the aqueous external environment into the larval zebrafish gut.

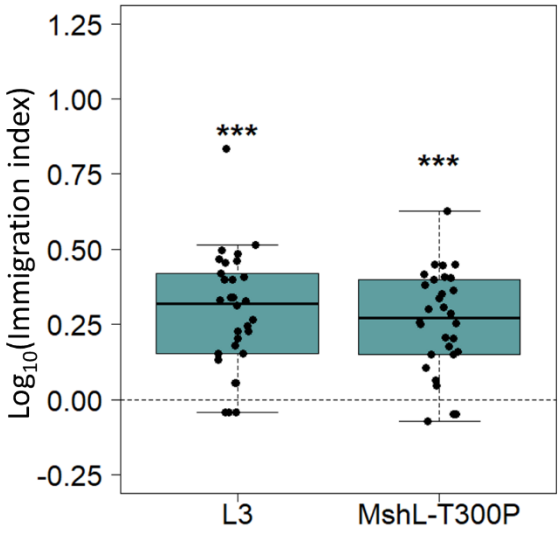

Figure 12: Relative per capita immigration of L3 and MshL-T300P mutant mutants compared to ancestor. Each box shows the per capita immigration efficiency of the indicated strain relative to the ancestor. Each point represents a single dissected and plated larval gut. The $\log_{10}$(immigration index) for each group of points was t-tested against a mu value of zero, my expected value if there were no competitive advantage, and Bonferroni corrected for multiple comparisons.

L3 and mshLT300P mutant are more motile:

Motility has been implicated in a number of host-microbe systems as being important for the maintenance of host-microbe symbioses (Raina et al., 2019; Stephens et al., 2015; Giraud et al., 2008; Van der Marel et al., 2008, Schlomann et al., 2018; Robinson et al., 2018), and recently, after employing a similar selection scheme to ours, Robinson and colleagues (2018) found that motility increased among evolving replicate lineages of a zebrafish-isolated *Aeromonas* species. In my system, enhanced motility could increase the efficacy of chemotaxis toward nearby hosts, thus increasing rate at which MR-1 cells encounter larval hosts, or it could aid in bypassing host filters (such as tight spaces or mucosal barriers), thereby increasing the rate of successful colonization when MR-1 cells come in contact with zebrafish larvae. Both proposed mechanisms would be consistent with the increased immigration efficiency I observed in my study.

Additionally, despite some ability to engage with surfaces, my biofilm results suggest that the L3 isolate and the MshL-T300P mutant are less adherent to physical surfaces and that this is likely due to a reduction in msh pili on the surface of MR-1 cells. A classical motif in microbiology is the trade-off between adherence and motility (Ferenci, 2016; Simm et al. 2004, Van Ditmarsch et al., 2013). Those cells that tend to adhere more also tend to swim less and vice versa. Given that my observation that evolved and mutant MR-1 cells were less adherent, and that previous research has highlighted motility as an important factor in host-microbe symbioses, I hypothesized that the L3 isolate and the MshL-T300P mutant would also be more motile. To assess this, I inoculated flasks containing GF larvae with a competition mixture of either L3a versus the MR-1 ancestor or MshL-T300P versus the MR-1 ancestor. In both cases, each competitor was tagged such that they expressed distinguishing fluorescent proteins. I then visualized the swimming speeds of each strain, and the proportion of each strain's population that was swimming, by combining an image-based cell tracking algorithm with fluorescence microscopy. I found that both the L3 isolate and the MshL-T300P mutant were more motile than ancestral strains used in this experiment. The adaptive strains were both faster swimmers (Figure 13, A and B), and they had a larger portion of their population that was motile (Figure 13C). These results suggest that motility is an important aspect of larval zebrafish colonization and highlight the role that the msh pilus plays in motility.

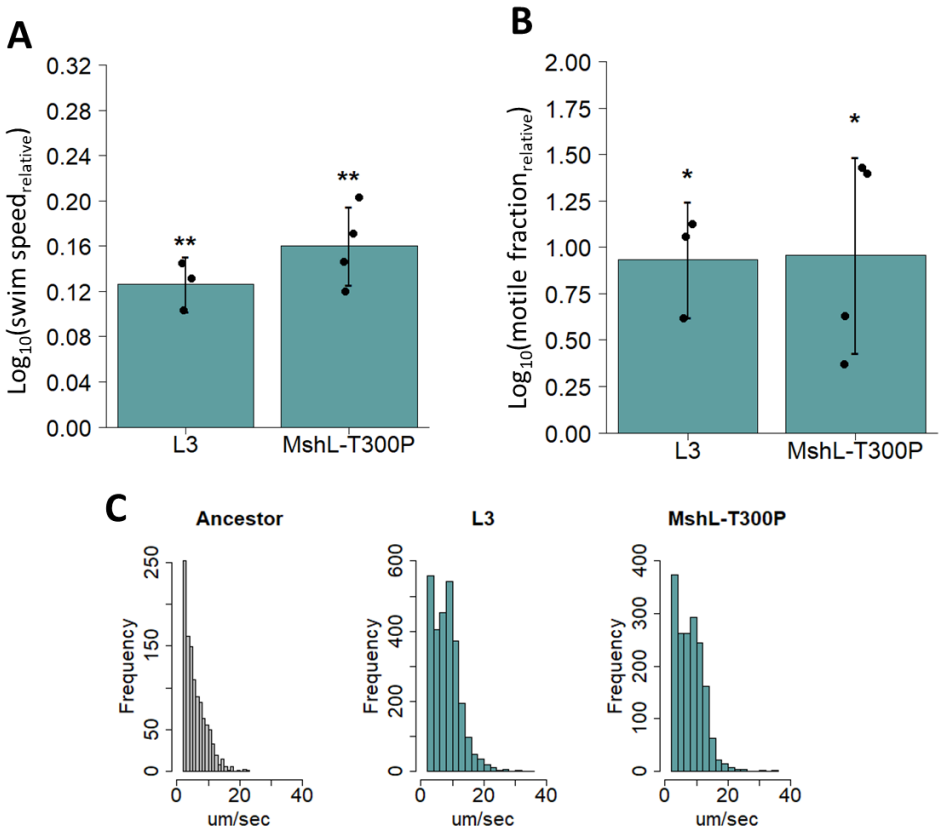

Figure 13: Motility characteristics of L3 and MshL-T300P mutant compared to ancestor. A) Mean swim speed of L3 and MshL-T300P relative to the ancestor. For each flask, represented by a point, the mean mutant swim speed was divided by the mean ancestral swim speed. B) Histograms of swim speeds for each strain are shown based on observations from three flasks for L3 (n = 2,740), four flasks for MshL-T300P (n = 1,713), and seven flasks for the ancestor (n = 1,104). Mutant strains have distributions that skew farther to the right implying faster swimming. C) The fraction of L3 and MshL-T300P cells that are motile divided by the fraction of ancestral cells that are motile. Each point represents the relative motile fraction for a single flask. Each flask contains data for either L3 or MshL-T300P and an ancestral strain. For A) and C) groups were t-tested against a mu value of zero (my null expectation) and Bonferroni corrected for multiple comparisons ($p<0.05$: *, $p<0.01$: **).

Discussion:

Here, I present the first known evidence describing the evolutionary trajectory of a non-host-associated bacterium as it adapts to a novel vertebrate host. I observed that an isolate that had been serially exposed to a zebrafish gut evolved to increase its representation in the planktonic portion of my experimental system. Additionally, I found

that adaptive strains had a higher per capita immigration rate than their unpassaged ancestor, suggesting an improved ability to bypass host filters to colonize larval guts. Both of these traits coincided with a mutation in an msh pilus operon that decreased biofilm formation and increased motility. Together this suite of evolved traits allowed an evolved MR-1 strain to outcompete its unpassaged ancestor in terms of its ability to colonize BF larval zebrafish digestive tracts.

As noted earlier, a classical paradigm in bacteriology is a trade-off between motility and biofilm formation (Ferenci, 2016). In keeping with this paradigm, the serial passage protocol employed in my experimental system selected on a mutation in the msh operon of MR-1 (MshL-T300P) which both decreased biofilm formation and increased motility. My characterizations of the effect of a loss of function mutation in this operon on MR-1 physiology and behavior are in line with what others have demonstrated for MR-1 (Saville et al., 2010; Bouhenni et al., 2010), and are consistent with *Vibrio cholerae* and *Aeromonas hydrophila* systems as well (Jones et al., 2015; Qin et al., 2014). Interestingly, the msh pilus has been implicated as an important modulator of host colonization in other host-microbe systems (List et al., 2018; Ariyakumar & Nishiguchi, 2009; O'Boyle et al., 2013), and some evidence suggests that this pilus interacts with the mammalian immune system (Hsiao et al., 2006; Hou et al., 2012). While the precise mechanism by which MshL-T300P increases larval gut colonization in my study remains elusive, it is likely that this mutation aides in the ability of the evolved L3 isolate to increase its planktonic representation. Unlike msh-pilus-expressing cells, which may more readily stick to inanimate surfaces due to their role in biofilm formation, MshL-T300P mutants form reduced biofilms and had increased motility which could result in

their exit back into the water column after an encounter with an inanimate surface. This proposed model agrees well with what others have observed in *V. cholerae* (Jones et al., 2015). In turn, since larvae are also present in the water column, having a more planktonic existence should increase the encounter rate between larval zebrafish and their potential MshL-T300P mutant colonists. Such dynamics would be consistent with the CIs I observe in my system-level, as well as LCM fitness assessments.

While outnumbering a potential competitor in the environment surrounding a host should provide a colonization advantage, it does not explain why I observed more efficient per capita immigration. In my migration experiment, each competitor was cultured separately in LCM prior to the start of the experiment. Because competition mixtures were then created by combining both cultures into a new container, all cells from both competitors should have been planktonic, rather than adherent to the side of the competition vessel which would have made them less accessible to larvae. Further, given the short timeframe I allowed for colonization in this experiment, it is unlikely the ancestor would have had time to settle out of the planktonic phase to any measurable degree. Therefore, the most parsimonious explanations for the differences I observed between each competing strain in the gut, was that the MshL-T300P-containing strains were either more efficient at bypassing host filters to colonize the gut, or they had evolved some method of increasing their relative host encounter rate. It is tempting to ascribe credit for an increased host encounter rate to the faster swimming speeds I observed for MshL-T300P strains, however it is not immediately clear how this would work. Anecdotally, when larval zebrafish are left undisturbed, they swim in short bursts, but spend a large portion of their time motionless. Specifically, during these times of

inactivity, because MR-1 cells are able to chemotax (Li et al., 2010), having a faster swimming speed could increase the efficacy of their chemotaxis, thus increasing their ability to encounter, and thus colonize, larval hosts (Raina et al., 2019). Alternatively, once a MR-1 cell encounters a host, faster swimming speeds could help MshL-T300P strains traverse narrow junctions within host tissues to reach the gut more quickly than their ancestral competitor (Raina et al., 2019). However, this mode of fitness enhancement, as a result of motility, may be less likely based on prior work that that showed no fitness advantage for a hypermotile *Aeromonas* species that was gavaged into the oral cavity of larval zebrafish (Robinson et al., 2018). Generating a mechanistic understanding of how increased motility might have influenced immigration efficiency in my system will require further research.

Interestingly, from a phenotypic perspective, my findings bear a striking resemblance to previous work conducted by Robinson et al. (2018) which employed a similar selection scheme. They experimentally evolved a zebrafish commensal species of *Aeromonas* (ZOR0001) in the presence of BF larval zebrafish to explore the ecological parameters that are relevant to host colonization and uncover which bacterial traits strengthen its pre-established association with this vertebrate host. They too found that increased relative fitness was associated with more efficient immigration and augmented motility. While my results bolster the generality of their findings, I had several reasons to expect different outcomes. In the present study I asked how a free-living species, with no record of host association, might evolve a novel host association. Given that Robinson and colleagues used a species that had been isolated from larval zebrafish, I speculated that this bacterium had evolved traits that were optimized for life among this aquatic host.

In keeping with this reasoning, the fact that this host isolate made the largest adaptive gains in traits related to the process of colonization, rather than growth inside the gut, may reflect the percentage of time this strain had spent evolving inside or outside the gut. If most of this host isolate's evolutionary history had taken place inside the gut, it may be well suited for life inside a gut, leaving the greatest room for improvement in traits related to the external portion of its host-associated life-cycle. In this way starting with a host isolate could have biased which selective pressures exert the strongest effect on *Aeromonas* adaptation.

Additionally, in other study systems aimed at assessing the role that biofilm formation plays in promoting the diversification of evolving lineages, two groups of researchers employed identical experimental evolution protocols on separate bacterial species. The results from these studies showed stark differences in the patterns of diversification each group observed in *Burkholderia cenocepacia* compared to *Pseudomonas aeruginosa* (Traverse et al., 2013; Flynn et al., 2016). Notably, the selection protocol used in these studies shares many similarities with the one featured here, and by Robinson and colleagues (2018). To be propagated in their serial passage protocol, cultured bacterial populations had to colonize plastic beads and grow in biofilms on the surface of those beads. Colonized beads were then transferred to new culture tubes containing sterile beads, and propagated strains would have to colonize these new beads to make it to the next passage. In our case, rather than beads, we used BF larval zebrafish. Given the differences in diversification patterns observed in the biofilm studies, and the different evolutionary histories of ZOR0001 and MR-1 used in the zebrafish studies featured here and in Robinson et al. 2018, I was surprised at the

remarkable levels of phenotypic convergence exhibited by ZOR0001 and MR-1. The vertebrate gut presents a diverse set of niches that provide fodder for adaptation (Donaldson et al., 2016). Since MR-1 had no known relationship with hosts, but had initially been isolated from an aquatic environment (Myers & Nealson, 1988; Venkateswaran et al., 1999), I expected that mutations with the largest adaptive value would have pertained to traits involved with *in vivo* fitness. That Robinson et al. (2018) and I observed convergent phenotypes in my studies, indicates the importance of transmission traits, and motility in particular, in adapting to and maintaining associations with a metacommunity of aquatic hosts. These results are consistent with previous work that has shown motility to be an important colonization factor of both zebrafish and other aquatic hosts (Raina et al., 2019; Stephens et al., 2015; Brennan et al., 2013; Van der Marel 2008).

Thinking beyond aquatic hosts, there are several aspects of my findings that may be generalizable. First my msh pilus operon data add to the growing body of evidence that implicates this operon in biofilm formation, motility, and indeed host association. Second, as indicated by the LCM competition data, I have provided evidence that novel host symbionts can use the environment that hosts construct as an adaptive stepping stone to become host-associated. Thus, as future researchers examine which bacterial species might evolve the potential to impact host health, it will be important for them to consider how host habitats select for traits that could encourage host association. Finally, it is clear from my work and others', that evolving traits that increase transmission into and between groups hosts is vital to the initiation and maintenance of host-microbe symbioses. To disentangle the effect that selection by groups of cohoused hosts has on an

evolving bacterial lineage, it would be interesting to compare adaptive strategies pursued by lineages adapting to solitary hosts versus those adapting to groups of cohoused hosts. Such an experiment might have a higher likelihood of selecting traits that increase *in vivo* fitness.

BRIDGE:

Many evolutionary biologists have wondered, for any given environment, how many adaptive pathways are available to an evolving organism. Researchers have attempted to understand this by evolving replicate populations of highly reproductive organisms under controlled circumstances, and such approaches have illustrated that even when genetic variability is low in a founding population, multiple distinct genotypes can give rise to elevated fitness (Kvitek & Sherlock, 2011; Khan et al., 2011). In some cases, these genotypes are associated with convergent phenotypes (Tenaillon et al., 2012; Fong et al., 2005), implying the existence of a limited number of adaptive trajectories, while in other cases divergent phenotypes have been selected (Rainey & Travisano, 1998). In Chapter II I found that in my replicate evolving lines of MR-1, selection commonly favored mutations in a msh pilus operon, however these were not the only mutations that accumulated during my serial passage experiment. Given the extent of niche space available in my experimental system, and in host-based environments in general, I hypothesized that there could be multiple adaptive pathways available to bacterial lineages once they encounter a host. To address this question, in Chapter III I compare the phenotypes of two evolved isolates from different replicate populations that have unique sets of mutations. Specifically, I focus on the same isolate I examined in Chapter II, and I contrast its behavior and fitness with an evolved isolate from a different population containing an adaptive mutation in a gene with potential diguanylate cyclase and phosphodiesterase activities.

CHAPTER III:

SELECTION ON TWO SEPARATE GENETIC PATHWAYS INCREASES MOTILITY AND GUT COLONIZATION IN LARVAL ZEBRAFISH

Introduction:

Given their long history on this planet (Hug et al., 2016), and their propensity for rapid adaptation (Van den Bergh, et al., 2018), bacterial lineages have radiated into practically every imaginable niche (Alivisatos et al., 2015; Thompson et al., 2017). Because animals contain a wide breadth of niche space and represent a significant and reliable source of nutrients, bacteria have routinely formed novel symbioses with animal hosts (McFall-Ngai et al., 2013). However, in order to avail themselves of host derived resources, bacterial lineages must survive a barrage of diverse selective forces (Bakke et al., 2015; Ley et al., 2008; Donaldson et al., 2016; Sass et al., 2010; McLoughlin et al., 2016; Quinn et al., 2018; Friedman et al., 2018; Ottman et al., 2017; Segal, 2005). Consequently, due to the enormity of phenotypic variation that exists across all bacteria, and the fact that organismal evolution is often highly constrained by trade-offs in adaptive strategies, most bacteria are not well suited for life in association with an animal host. For the comparatively few species that are capable of colonizing animal hosts, it is unknown how bacteria evolve suites of traits that enable this ability, or whether there are a limited number of evolutionary trajectories that are likely to result in host association.

As I discussed in Chapter II, experimental-evolution-based approaches have been fruitful in elucidating how bacterial populations evolve and adapt under a number of different contexts (Lenski, 2017; Van den Bergh et al., 2018). Because of evolutionary dynamics such as clonal interference, individual bacterial lineages within an evolving population must rely heavily on the emergence of novel mutations to access new

beneficial traits, and then lineages containing such traits compete for resources (Barrick & Lenski, 2013). If resource requirements between two evolving lineages overlap, the lineage with greater fitness will tend to increase its frequency in the population at the expense of the other (Gerrish & Lenski, 1998). This can result in reduced diversity if a lineage containing a beneficial trait sweeps to fixation (Cira et al. 2018). Alternatively, if an environment contains multiple types of resources, adaptive mutations can allow evolving lineages to exploit each resource separately (Barrett & Bell, 2006, Herron & Doebeli, 2013). If such mutations impart a similar adaptive value to the lineages containing them, or adaptation to one resource precludes the exploitation of the other, lineages can coexist (Spencer et al., 2008; Frenkel et al., 2015). Additionally, environmental structure can either help to limit competition between evolving lineages by restricting physical interaction between lineages, or by partitioning resources such that diverging lineages can avoid niche overlap (Rainey & Travisano, 1998; Bailey & Kassen, 2012). Even in relatively constant environments with reduced niche space, new niches can arise through mechanisms such as cross-feeding, whereby one strain gains the ability to metabolize the byproducts of another (Turner et al., 1996).

In a model vertebrate system such as ours, the existence of a diverse set of selective pressures would imply a variety of ways to adapt to the vertebrate gut environment. Despite this, experimental investigations in other vertebrate systems have found high degrees of evolutionary parallelism when examining the evolution of bacteria adapted to the vertebrate gut in that, within each study, the same genomic region or phenotype was targeted by selection (Barroso-Batista et al., 2014; Giraud et al., 2008; Lescat et al., 2017; Robinson et al., 2018; Zhao et al., 2018). However, the bacterial

species examined in each of these studies encompass lineages that have pre-established evolutionary histories with their respective vertebrate hosts. Thus, the evolutionary patterns observed in these studies may be the result of fine-tuning genomes which are already optimized for life in a vertebrate gut. Further, many evolutionary events can occur as a free-living bacterium transitions to life as a host associate, and during these transitions, intermediate forms can be lost, thereby precluding my understanding of how the transition occurred.

I aimed to capture such a transition in its early stages by serially exposing replicate populations of a bacterial species that has no prior known history with an animal host to the digestive tract of a model vertebrate. On one hand, given the ample breadth of niche space available to an evolving population of *Shewanella oneidensis* (MR-1) cells adapting to groups of bacteria free (BF) larval zebrafish, both within the gut and externally in the aqueous environment, I hypothesized that naive MR-1 populations would pursue divergent adaptive trajectories resulting in unique phenotypes that would allow for the exploitation of distinct niches. Alternatively, if the strongest selective pressures in my experiment encourage adaptation to a single facet of the whole system, I expected that clones which specialize in the exploitation of that single niche would exhibit high relative fitness, leading to phenotypic convergence in competing lineages. To distinguish between these hypotheses, I examined the phenotypes of two classes of adaptive mutations that accumulated during my serial passage scheme. Despite the diversity of niches available in my system, I found elevated levels of phenotypic overlap between divergent genotypes. Both a mutation in a mannose sensitive hemagglutinin (msh) pilus gene and a mutation in a putative diguanylate cyclase (DGC) domain-

containing gene (locus tag: SO1551) exhibited elevated levels of motility. This augmented motility was associated with a more planktonic existence and increased levels of immigration into larval guts.

Methods:

Zebrafish husbandry:

To ensure animal specimens were treated ethically in all experiments involving zebrafish, I adhered to the standard protocols and procedures approved by the University of Oregon Institutional Animal Care and Use Committee (IACUC protocol: 15-98). BF derivations were carried out as described in Melancon et al., 2017. Details about larval gut dissections can be found the Serial passage section.

Bacterial strains:

My ancestral reference *S. oneidensis* (MR-1) and Shew-Z12 strains were obtained from Karen Guillemin's laboratory at the University of Oregon. All modifications to MR-1, including Tn7-mediated gfp and dTomato insertions as well as allelic exchange applications, were carried out as described in Wiles et al., 2018. Whole gene deletion allelic constructs were created via a splice by overlap extension protocol (Wiles et al., 2018), whereas evolved mutation allelic constructs were created via amplification of mutated segments of evolved genomes. All *S. oneidensis* strains were cultured in TSB at 30 °C under shaking conditions unless otherwise specified.

Serial passage:

My experimental evolution serial passage scheme was similar to the experimental design of Robinson et al. 2018 and was described in detail in my Chapter II methods section. Briefly, six replicate ancestral populations were then generated by combining subcultures of MR-1gfp (tagged using a Tn7-mediated insertion of a green fluorescent protein) and MR-1dT (tagged using a Tn7-mediated insertion of a dTomato fluorescent protein) at a 1:1 ratio. Each of these replicate ancestral populations were used to inoculate larval flasks containing ~15 mL of EM and ~15 BF larval zebrafish at 4 days post fertilization (dpf). Larvae were then incubated with MR-1 populations at 28 ˚C for 72 hours. At 7 dpf, 10 larvae were euthanized with tricane, mounted on a glass, and their digestive tracts were dissected out. Following dissections, all 10 digestive tracts culled from each flask were immediately homogenized using a bullet blender tissue homogenizer (Next Advance, Averill Park, NY, US) for 60 seconds at power 4. To preserve my ability to revive replicate populations after each passage, I created freezer stocks (stored at -80) by using a pipette to mix 200 µL from each homogenized tube with 200 µL of 50 % glycerol (25 % glycerol final concentration). The remaining contents of each homogenized tube was then stored at 4 ˚C for 0-14 days, at which point 250 µL were sampled to inoculate a subsequent set of BF larval flasks (~15 larvae in ~15 mL EM). This cycle was repeated for 20 passages. All six replicate evolving populations were maintained separately throughout my experiment.

Whole experimental system competitions:

5 mL overnight TSB cultures of competing strains were diluted 1:100 in TSB and allowed to grow out to late log stage (4-5 hours). 500 μL of each competitor was then mixed in a single 1.7 mL tube so that competitors were at an approximate 1:1 ratio. Competition mixtures were pelleted (7000 rcf for 5 min) and resuspended in 1 mL sterile EM. Resuspended competition mixtures were diluted 1:100, and 7.5 μL of these dilutions were used to inoculate BF larval flasks containing ~15 mL of EM and ~15 larvae at 4 dpf. Immediately following inoculation, triplicate 100 μL samples from each competition flask were dilution plated to establish the inoculation ratio of competitor1:competitor2 (CFU/mL). After inoculation, I incubated larval flasks at 28 ˚C for 72 hours. At 7 dpf, multiple larvae per flask were euthanized with tricane, and their guts were dissected and individually placed in 1.7 mL tubes containing 500 μL of sterile EM and ~100 μL 0.5 mm zirconium oxide beads (Next Advance, Averill Park, NY, US). The contents the larval guts in each these tubes were then immediately homogenized using a bullet blender tissue homogenizer (Next Advance, Averill Park, NY, US) for 60 seconds at power 4. Homogenized tubes were then dilution plated to determine the CFU/gut for each competitor. A competitive index (CI) was calculated for each dissected gut by dividing the ratio of competitor1:competitor2 found in each gut by the mean inoculation ratio determined from the triplicate measurements for each corresponding flask, CI = $\left(\frac{Competitor1{:}Competitor2_{guts}}{Competitor1{:}Competitor2_{inoculum}}\right)$.

Biofilm assay:

Triplicate biological replicate overnight TSB cultures of strains of interest were diluted 1:100 in TSB and allowed to grow out to mid-late log stage (3-4 hours). Optical densities were measured at 600 nm for each overnight culture. 1 mL of subcultured strains were pelleted (7000 rcf for 5 min), and the pellets were resuspended with a volume of sterile EM that would correspond with a TSB optical density of 1.0 if read at 600 nm. 150 μL of each resuspension were added to four wells of a round-bottom 96 well polystyrene plate (Greiner Bio-One, catalogue #: 650185) per biological replicate. The plate was incubated at 30 ˚C for 24 hours, and the volume of each well was removed with a multichannel pipette. The wells were rinsed twice with 200 μL of sterile EM, and each well was stained with 180 μL 0.1 % crystal violet (CV). The plate was incubated at room temperature for 10 min, at which point the crystal violet was removed with a multichannel pipette, and the wells were again rinsed twice with 200 μL of sterile EM. The CV was solubilized with 100 % dimethyl sulfoxide (DMSO) for 15 min with agitation (~180 rpm on a rotating minishaker), and 100 μL of the solubilized CV was added to 100 μL of 100 % DMSO in a flat-bottom 96 well polystyrene plate (Corning Incorporated, reference #: 3595). Optical densities were then measured for each well at 570 nm.

Motility assays:

5 mL overnight TSB cultures of strains of interest were diluted 1:100 in TSB and allowed to grow out to late log stage (4-5 hours). Strains were then prepped for inoculation as described in my whole experimental system competitions or my colonization level

assessments above (pellet, resuspend in sterile EM, and dilute) depending on whether I was assaying motility under competitive or monoassociation conditions respectively. 7.5 μL of each strain prepared for inoculation were used to inoculate BF larval flasks containing ~15 mL of EM and ~15 larvae at 4 dpf. Inoculated larval flasks were incubated at 28 ˚C for 13-17 hours, at which point bacteria in each flask were imaged on an inverted microscope (Nikon Eclipse Ti-e) by focusing on the bottom interior surface of the flask. Ten, 30 second movies were then captured at a rate of > 15 frames per second. When movies were taken of competing populations, each population was fluorescently tagged with either gfp or dTomato, and movies were taken separately to capture the motility dynamics of each tagged population independently within the same flask. When I examined flasks containing only a single population (monoassociation), movies were taken in bright field. To calculate motility characteristics such as the speed of motile cells and the proportion of a population that was motile, I relied on a particle tracking algorithm (Parthasarathy, 2012; source code: http://pages.uoregon.edu/raghu/particle_tracking.html). Only cells deemed to be moving faster than 2 um/second were considered motile.

Comparative genomics:

I submitted my MR-1 and Shew-Z12 strains to the Washington State University Molecular Biology and Genomics Core (WSUGC) for long read sequencing. Genome assembly for MR-1 was conducted by WSUGC, whereas genome assembly for Shew-Z12 was conducted in house with Canu v. 1.7.1 (Koren et al., 2017). To generate

annotation files for these genomes, I relied on Prokka v1.12 (Seemann, 2014), and RAST v. 2.0 (Aziz et al., 2008).

Genome comparisons between S. oneidensis and non-S. oneidensis species:

To assess the diguanylate cyclase and phosphodiesterase content in *S. oneidensis* and Shew-Z12, I searched the publicly available genomes of *S. oneidensis* and Shew-Z12 curated by Integrated Microbial Genomes and Microbiomes (IMG/M, https://img.jgi.doe.gov/; Chen et al., 2018) for all genes containing protein family (pfam) terms that corresponded with GGDEF and EAL domains (pfam00990 and pfam00563 respectively; See Appendix Table 3 for metadata).

Specific gene and operon comparisons between S. oneidensis and Shew-Z12:

Comparisons between MR-1_SO1551 and Shew-Z12_ L976_03566 were visualized by creating Clustal Omega-based multiple sequence alignment files that were imported into Jalview2 (Waterhouse et al., 2009). To highlight regions of divergence between genes I color-coded my comparisons using the color by annotation feature of Jalview2 (Figure 5; Figure 17). This feature color-codes amino acid comparisons per site based on biochemical conservation.

Evolved mutation calling:

I determined evolved mutations as described in the "Evolved mutation calling" portion of Chapter II's "Comparative genomics" methods section. Briefly, I sequenced the genomes of

4 isolates from evolved replicate populations 2, 3, and 5, and ancestral strains MR-1gfp and MR-1dT. I then compared these genomes to my wild type ancestral MR-1 reference strain using breseq 0.31.0 in consensus mode, and looked for single nucleotide

polymorphisms (SNPs) and indels that were present in each evolved isolate, but absent in MR-1gfp and MR-1dT isolates. Gene annotations for the mutations listed in Table 1 were determined by Prokka v1.12 (Seeman, 2014). The mshOP gene annotations were determined by RAST v 2.0.

Colonization level assessments:

5 mL overnight TSB cultures of strains of interest were diluted 1:100 in TSB and allowed to grow out to late log stage (4-5 hours). 1 mL of subcultured strains were then pelleted (7000 rcf for 5 min), and the pellets were resuspended in 1 mL sterile EM. This resuspension was then diluted 1:100, and 7.5 μL of this dilution was used to inoculate BF larval flasks containing ~15 mL of EM and ~15 larvae at 4 dpf. After inoculation, I incubated larval flasks at 28 ˚C for 72 hours. At 7 dpf, multiple larvae per flask were euthanized with tricane, and their guts were dissected and individually placed in 1.7 mL tubes containing 500 μL of sterile EM and ~100 μL 0.5 mm zirconium oxide beads (Next Advance, Averill Park, NY, US). The contents the larval guts in each these tubes were then immediately homogenized using a bullet blender tissue homogenizer (Next Advance, Averill Park, NY, US) for 60 seconds at power 4. Homogenized tubes were dilution plated to determine the CFU/gut.

Immigration and in vivo growth assays:

Separate 5 mL overnight TSB cultures of competitor strains of interest were diluted 1:100 in TSB and allowed to grow out to late log stage (4-5 hours). 1 mL of subcultured strains were then pelleted (7000 rcf for 5 min), and the pellets were resuspended in 1 mL sterile

EM. Resuspensions for each competitor were then diluted 1:100, and 7.5 μL of each dilution was used to inoculate separate ~15 mL LCM flasks (created as described in LCM competitions assays above). Each competitor was incubated in its own LCM flask at 28 ˚C for 12-15 hours at which point two flasks (one containing each competitor) were combined into a single polystyrene petri dish such that the petri dish contained a ~30 mL competition mixture. 30-35 4 dpf larvae were then added to each competition mixture dish, and 100 μL samples were immediately taken in triplicate from each dish. These samples were dilution plated to establish an inoculating competition mixture (competitor1:competitor2). After a 40-60 min incubation period at 28 ˚C, the LCM was removed and replaced with 100 mL sterile EM for a total of three rinses. These media exchanges were conducted to reduce the density of each *S. oneidensis* competitor in the EM so that both populations were below a threshold where colonization would readily occur (colonization threshold > 10^3 CFU/mL; Figure 8A & 8B). 10 larval guts were then dissected as described above in my whole experimental system competitions. By dilution plating the contents of each dissected gut, I established a competition ratio for each gut and determined a mean competition ratio for a typical larval gut (competitor1:competitor2). Following gut dissections, EM exchanges were repeated every two hours to ensure that bacterial loads in the EM remained below 10^3 CFU/mL (monitored by plating a single 100 μL sample just prior to the beginning of each set of EM exchanges). The larvae that remained in the competition petri dishes after the first set gut dissections were completed, were incubated at 28 ˚C between sterile EM exchanges. 8.5-10.5 hours after initial gut dissections took place, 10 additional larval guts were dissected, and their contents dilution plated to establish a final mean competition ratio for

a typical larval gut (Competitor1:Competitor2). Since the CFU/mL in the EM was maintained at a low level between bouts of dissection, any increase in the mean CFU/gut I observed between dissection bouts should have been affected primarily by the growth of *S. oneidensis* populations *in vivo*. Using the initial mean CFU/gut and the final mean CFU/gut that I observed for each competing strain, I then calculated a per capita growth rate for each strain per competition ($r = \ln\left(\frac{CFU/gut_{final}}{CFU/gut_{initial}}\right)$). I then used my per capita growth rate to calculate an immigration index and relative *in vivo* fitness metric (immigration index: $\frac{gut\ ratio_{competitor1:competitor2}}{inoculation\ ratio_{competitor1:competitor2}} * e^{(r_{competitor2} - r_{competitor1})*t}$; relative *in vivo* fitness: $\frac{r_{competitor1}}{r_{competitor2}}$). In this way, my *in vivo* relative fitness compares the relative per capita growth rates I observed *in vivo* for each competitor, and I are able to use these per capita growth rates to control for any growth that may have occurred *in vivo* during my immigration assay.

Competition model:

To model two bacterial strains whose populations stochastically collapse while competing against each other in a gut, I employed a discrete Lotka-Volterra competition model that assumed each species competed for resources equally with each other (i.e. the competition coefficients for each species were set to a value of 1). I used my experimental data to approximate carrying capacity (~50,000 CFU/gut), growth rates (r_1 = 0.21, r_2 = 0.2), and founding populations (~200 CFU/gut per strain). I then initiated a stochastic variable that imparted a 10 % chance of population collapse at each time point, and population collapse proportions were determined by an additional stochastic variable.

Thus, for any given population collapse, the proportion of the population that remained would be between 0 % and 100 %, exclusive. I also assumed that when populations collapsed, both competitors would lose the same proportions of their populations.

Larvae-conditioned media (LCM) competitions:

Overnight TSB cultures of competing strains were diluted 1:100 in TSB and allowed to grow out to late log stage (4-5 hours). 500 μL of each competitor was then mixed in a single 1.7 mL tube so that competitors were at an approximate 1:1 ratio. Competition mixtures were pelleted (7000 rcf for 5 min) and resuspended in 1 mL sterile EM. Resuspended competition mixtures were diluted 1:100, and 7.5 μL of these dilutions were used to inoculate BF larval flasks where larvae had been removed at 4 dpf. Each of these BF larval flasks had contained ~15 mL of EM and ~15 larvae that had been conditioning their EM over the course of four days. Multiple BF larval flasks had been combined into a single sterile vessel, and the larvae were removed using a 1.5 mL bulb pipette, leaving behind only the LCM. ~15 mL of LCM was then aliquoted into fresh sterile polystyrene culture flasks. These flasks were inoculated as stated above and triplicate 100 μL LCM samples were taken per competition flask and dilution plated on TSA to establish an initial competition ratio (competitor1: competitor2). Each competition flask was incubated at 28 ˚C for three days (to mimic a single passage during my serial passage procedure) at which point 100 μL samples were again taken in triplicate and dilution plated to establish a final competition ratio (competitor1:competitor2). Competitive indices were calculated by dividing the final CFU ratio of competitor1:competitor2 by the inoculation ratio, CI = $\left(\frac{Competitor1{:}Competitor2_{final}}{Competitor1{:}Competitor2_{inoculum}}\right)$.

Results:

Several types of evolved mutations improve fitness:

In Chapter II, I described how after serially passaging replicate populations of the non-host-associated bacterium MR-1 through the digestive tracts of groups of BF larval zebrafish, I observed high degrees of evolutionary parallelism associated with selection on an msh pilus operon (mshOP). Specifically, by comparing the genome sequences of evolved isolates to the sequences of their unpassaged ancestor, I found that 20 out of 24 sequenced isolates from six replicate evolved populations (four isolates per population) contained mutations in the mshOP. However, given that four isolates contained no mutations in the mshOP, I wondered whether there were multiple adaptive pathways MR-1 populations could pursue to achieve higher relative fitness in larval zebrafish guts.

Three out of four of these isolates stem from evolved line 5 (L5), whereas the remaining isolate was found in line 2 (L2; Table 1). Interestingly, both L2 and L5 contained isolates that had mutations in the mshOP, however only one of four isolates in L2 were devoid of mshOP mutations while three of four isolates in L5 were devoid of mshOP mutations. Based on clonal interference dynamics, I would expect that the adaptive value of a given mutation should be reflected by its representation in an evolving population. Thus, my sequencing results imply that the non-mshOP mutations found in L2 were of inferior adaptive value to mshOP mutations (due to their lower representation), while the non-mshOP mutations found in L5 were of superior adaptive value (due to their higher representation). To test this, I compared the performance of non-mshOP-mutation-containing isolates from L5 (isolate L5b) and L2 (isolate L2a) to that of my best performing mshOP-mutation-containing isolate from line 3 (L3a) in terms

of their competitive ability against their unpassaged ancestor (Note: L3a is the same isolate I examined in Chapter II). Consistent with my expectations, I found that L5b had the highest competitive fitness (Figure 14; see methods for a description of how competitive indices were calculated). While L3a did not statistically outperform L2a, this comparison yielded a p value of 0.087, and the mean of the log-transformed competitive indices for L3a was ~79% higher than that of L2a (l2a = 0.63 $\pm$ 0.4, L3a = 1.13 $\pm$ 0.5; mean($\log_{10}$(CI) $\pm$ SD($\log_{10}$(CI))), suggesting that L3a may have slightly higher competitive fitness than L2a. Given that L2a seemed to have the lowest adaptive value of these three isolates, I chose not to examine L2a further, and to continue comparing L3a and L5b for subsequent analyses instead.

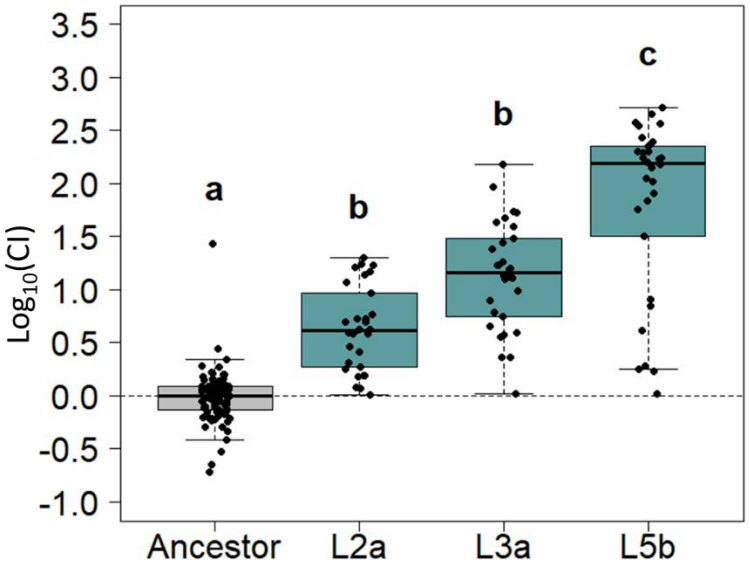

Figure 14: Competitive ability of MR-1 isolates from replicate evolved populations against MR-1wt. Each point represents the CI measured for a single larval gut. Ancestral competition against itself is shown as a control to represent the absence of a competitive advantage. Statistical groupings are indicated by letters above each box for a significance threshold of $p<0.05$. Letters in common between groups indicate the absence of a significant difference in each group's mean.

L5b's fitness stems from a mutation in a gene containing diguanylate cyclase and phosphodiesterase domains:

Given that L5b did not contain any mshOP mutations (Table 1), I wondered which genetic changes might be responsible for its increased fitness. Because L5b had no mutations in common with isolates from any other evolved population (Table 1), I was unable to determine which of the four mutations found in L5b might improve its fitness strictly by relying on gene annotations alone. I therefore used allelic exchange to recreate each L5b mutation individually in the ancestral genomic background and tested their fitness effects via competition against the ancestor. The only mutation that individually had any effect on competitive fitness against the ancestor was a mutation that occurred in a gene containing putative phosphodiesterase (PDE) and diguanylate cyclase (DGC) domains (locus tag: SO1551; Figure 15). This mutation consisted of a single base pair deletion about a quarter of the way through SO1551, and thus it is likely a loss of function mutation. Although, the SO1551 mutation seemed to improve fitness to a lesser degree than that achieved by L5b, competing L5b against the SO1551 mutant dramatically reduced L5b's competitive index (Figure 15, right-most competition). This result suggests that the SO1551 mutation is capable of explaining the majority of L5b's improved fitness compared to the ancestor.

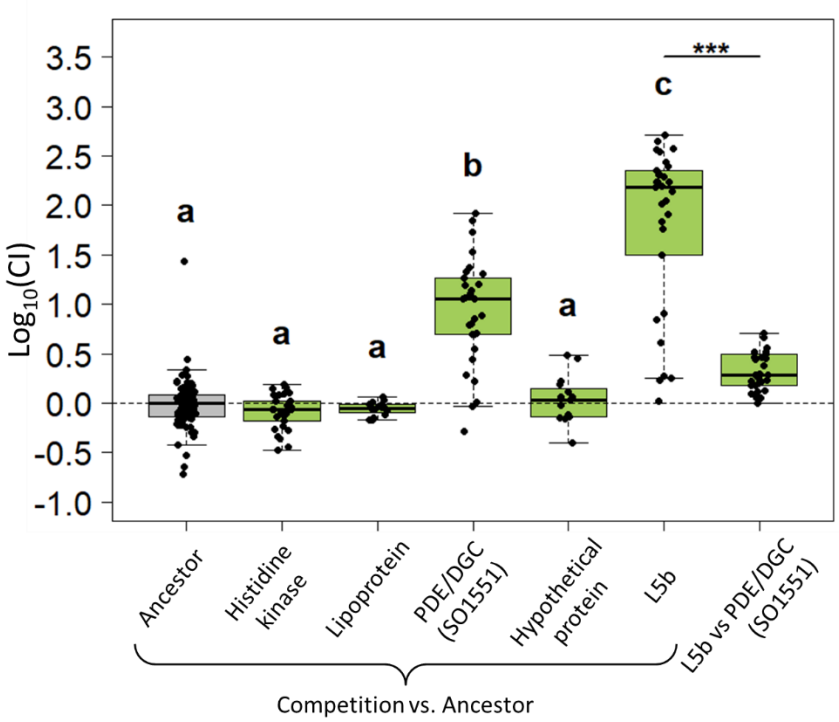

Figure 15: Competitive ability of L5b and its associated mutants against the MR-1 ancestor. Except for SO1551, which putatively contained both PDE and DGC domains, Prokka-based annotations were used to distinguish specific mutants along the X-axis. Each point represents the competitive index (CI) measured for a single larval gut. Ancestral competition against itself is shown as a control to represent the absence of a competitive advantage. Statistical groupings are indicated by letters above each box for a significance threshold of $p<0.05$. Letters in common between groups indicate the absence of a significant difference in each group's mean. Statistics for the L5b vs SO1551 mutant competition (right-most box) were conducted separately since this competition had unique competitors ($p<0.001$: ***).

L5b forms normal biofilms:

Previously, I had determined that the causative mutation underlying L3a's fitness improvement was a loss of function mutation in the mshL gene of the mshOP (MshL-T300P). After obtaining definitive evidence that L3a and L5b had adapted to groups of larval zebrafish via different genetic pathways, I sought to understand whether each genetic pathway was associated with unique sets of phenotypes. Numerous studies have implicated cyclic dimeric guanosine monophosphate (c-di-GMP) in a multitude of

cellular processes in bacteria, including biofilm formation (Römling et al., 2013; Jenal & Malone, 2006; Jones et al., 2015; Chao et al., 2013). This secondary messenger molecule is synthesized by DGCs and degraded by phosphodiesterases (PDEs) in a cyclic fashion (Simm et al., 2004), and in particular, the over expression of DGCs or reduced expression of PDEs has been shown to cause an increase in biofilm formation (Chao et al., 2013; Jones et al., 2015; Antoniani et al., 2010; Kirillina et al., 2004; Thormann et al., 2006). Interestingly, the one base pair deletion I observed in SO1551, happened downstream of an EAL motif, which is normally involved in PDE activity (Schmidt et al., 2005), and upstream of a canonical GG(D/E)EF motif (GGEEF in SO1551), which is normally involved in DGC activity (Merritt et al., 2010). This suggests that this deletion in SO1551 could disrupt DGC activity while leaving PDE activity intact. In keeping with the literature on DGCs and PDEs, if the mutation I observed in SO1551 was indeed a loss of function mutation, it is possible this mutation could cause a reduction in MR-1' biofilm forming capacity. Further, since my serial passage scheme produced adaptive isolates that had mutations in putative PDE/DGC-domain-containing genes (L2a and L5b) as well as the mshOP (L3a), and my prior work (Chapter II) documented a loss of function mutation in the mshOP as having a negative impact on biofilm formation, I hypothesized that the SO1551 mutation would result in L5b forming reduced biofilms compared to the ancestral MR-1 strain. I therefore compared L5b's biofilm phenotype to that of L3a and the wild type (WT) ancestor in larvae-conditioned medium (LCM). LCM was created by deriving zebrafish embryos BF, allowing them to develop into larvae over 96 hours (the typical time frame when I inoculate larvae during competition assays), and then removing the larvae from their embryonic medium (EM), leaving behind LCM. I found L5b's

biofilm phenotype was comparable to that of the ancestor (Figure 16). Since msh pilus expression has been associated with the formation of biofilms (Saville et al., 2010; Jones et al., 2015), these results suggest that the SO1551 may not alter the expression of the msh pilus.

I also included the *Shewanella* zebrafish isolate (Shew-Z12) in this assay to see if either L3a or L5b shared a common phenotype with this closely-related *Shewanella* species. Surprisingly, the Shew-Z12 isolate had a phenotype that was statistically indistinguishable from the wt ancestor (Figure 16). Moreover, given that L3a, L5b, and Z12 are all capable of outcompeting the ancestor in terms of their ability to colonize BF larval zebrafish guts, and yet these strains display a range of biofilm phenotypes under the conditions tested, it appears that my biofilm assay is not capable of predicting fitness. One reason for this could be, that my biofilm assay involves suspending cell cultures in LCM at a density that far exceeds their carrying capacity in this medium. Thus, my results may not be representative of how these strains behave during a normal fitness competition.

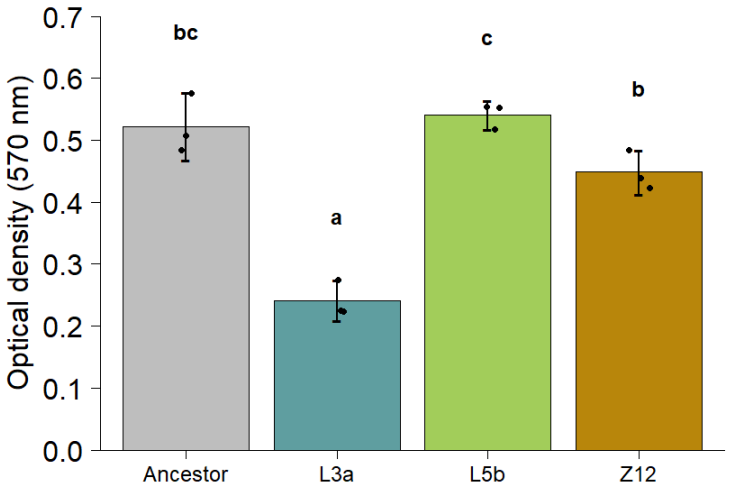

Figure 16: Crystal violet biofilm assay comparing L3a, L5b, and *Shewanella sp. ZOR0012* to ancestor. Optical density (570 nm) corresponds to crystal violet intensity. Higher optical density readings indicate more robust biofilms. Statistical groupings are indicated by letters above each box for a significance threshold of $p<0.05$. Letters in common between groups indicate the absence of a significant difference in each group's mean. Only L3a (containing MshL-T300P) shows a reduction in biofilm formation compared to the wt ancestor.

L3a and L5b show similar enhancements in motility, consistent with the Behavior of Shew-Z12:

Both DGCs and PDEs have also been shown to modulate bacterial motility (Wei et al., 2016; Liu eta l., 2010), and several prior studies have highlighted traits related to bacterial transmission and dispersal, including motility, as having a role in maintaining host associations (Robinson et al., 2018; Raina et al., 2019; Stephens et al., 2015; Van der Marel et al., 2008). Additionally, a commonly observed trade-off in microbiology is that it is difficult for bacteria to be simultaneously motile and prone to biofilm formation (Simm et al. 2004, Van Ditmarsch et al., 2013). Consistent with these observations, in Chapter II I observed that the evolved L3a isolate was more motile than the ancestor, under competitive conditions, and less able to form biofilms. Since L5b showed a biofilm forming capacity that was equivalent to the WT ancestor, I expected that L5b might have

reduced relative motility characteristics compared to L3a. To test this hypothesis, I determined the swimming speeds and motile proportion of L5b's population in a flask containing larval zebrafish. These assays were conducted in competition with the ancestor so that both L5b and the ancestor could be assessed under identical conditions. Each competition flask was placed under an inverted microscope, and I then generated movies of MR-1 cells swimming near the bottom of a larval flask. A tracking algorithm was used to calculate relevant metrics (Parthasarathy, 2012). Surprisingly, given L5b's biofilm-forming capacity, I observed that both L3a and L5b appeared to have faster swim speeds and larger fractions of their populations that were motile compared to the ancestral strain (Figure 17, A and B). Further, with respect to their advantage over the ancestor, these evolved isolates appeared to perform similarly via the two motility metrics I tracked.

To see if the motility phenotypes I measured for L3a and L5b were in line with adaptive trajectory taken by the Z12 zebrafish isolate, I performed additional experiments comparing Z12 to the wt MR-1 ancestor. These experiments had to be conducted under monoassociation conditions (rather than competitive conditions), as previous attempts to fluorescently tag Z12 were unsuccessful, preventing us from distinguishing Z12 from MR-1 within the same flask. Across three replicates for each strain, I observed that Z12 too had augmented swimming speeds and motile fractions compared to the MR-1 ancestor (Figure 17, C and D). Because all three isolates I characterized were able to outcompete the MR-1 ancestor to colonize BF zebrafish larvae, the results from my analyses confirm motility as a key component of host association.

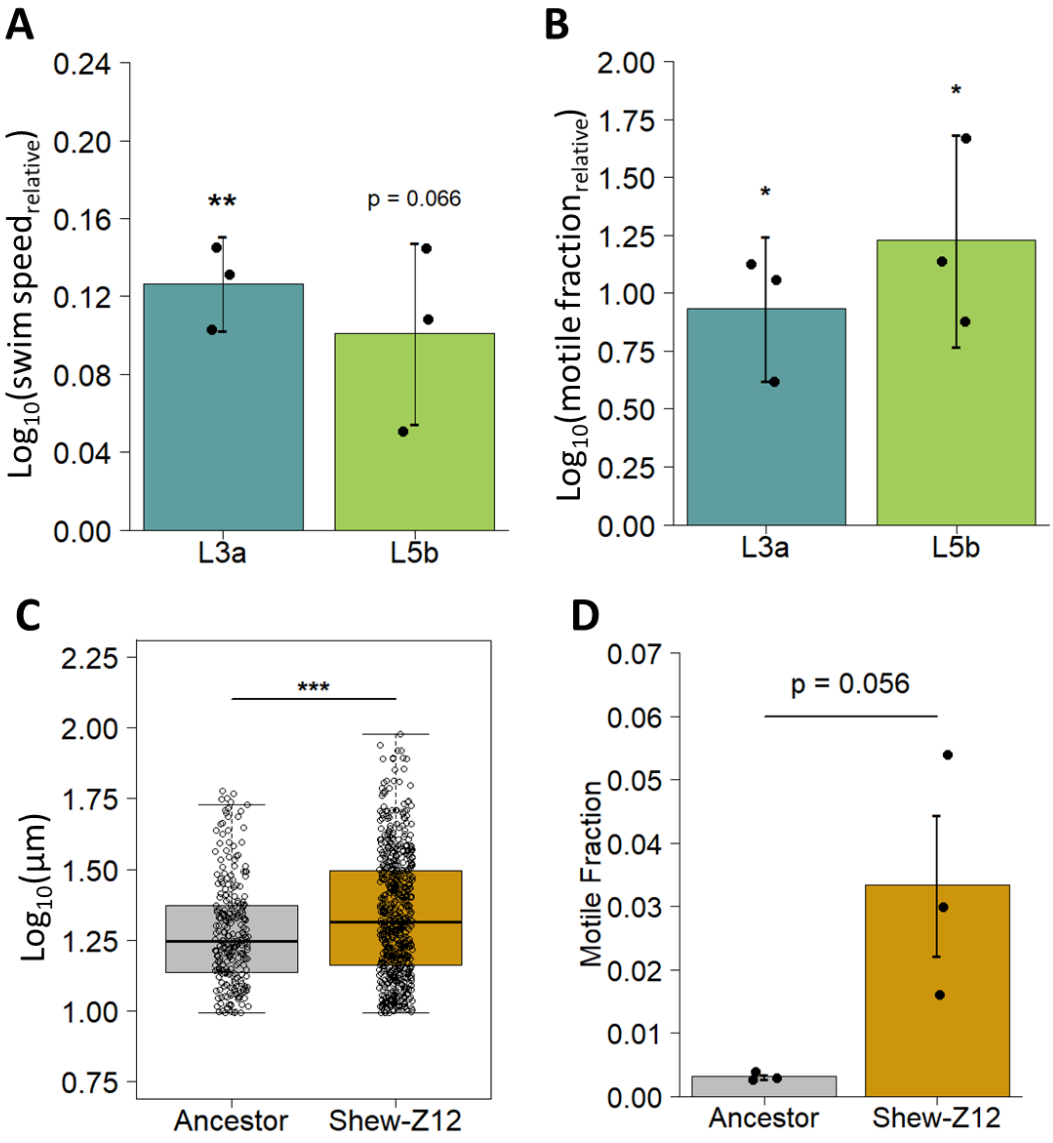

Figure 17: Motility characteristics of L3a, L5b, and *Shewanella sp. ZOR0012* compared to ancestor. A) Mean swim speed of L3a and L5b relative to the ancestor. For each flask, represented by a point, the mean mutant swim speed was divided by the mean ancestral swim speed. B) The fraction of L3a and L5b cells that are motile divided by the fraction of ancestral cells that are motile. Each point represents the relative motile fraction for a single flask. Each flask contains data for either L3a or L5b and an ancestral strain. For A) and B) groups were t-tested against a mu value of zero (my null expectation) and Bonferroni corrected for multiple comparisons ($p<0.05$: *, $p<0.01$: **). C) Swimming speeds observed for the MR-1wt and Shew-Z12. Each point is a single observation. A one-tailed t-test was conducted to determine whether Shew-Z12 had a significantly faster mean than MR-1wt ($p<0.001$: ***). D) The fraction of each population that is motile in three biological replicates for MR-1wt and Shew-Z12. Each point represents a mean for a single flask.

SO1551 is highly conserved in MR-1 and Shew-Z12:

Because L5b and Shew-Z12 were both able to outcompete my ancestral reference MR-1 strain, and L5b and Shew-Z12 performed similarly in my motility and biofilm assays, I wondered whether the mutation I observed in SO1551 represented evidence of evolutionary convergence between L5b and Shew-Z12. Both MR-1 and Shew-Z12 contain orthologues of the SO1551 gene (locus tag: L976_03566 in Shew-Z12) which are 99 % similar at the amino acid level (Figure 18) and have the same number of base pairs. Given the presence, high sequence identity, and identical length of these two genes, it is probable that both SO1551 and L976_03566 perform similar functions in their respective organisms. The fact that the one base pair deletion I observed in L5b's SO1551 allele (L5b-SO1551) happened downstream of a PDE domain (EAL) but produced a premature stop codon upstream of a DGC domain (GGEEF), could leave intact L5b-SO1551's PDE activity while negating its DGC activity. This makes it exceedingly unlikely that the evolved allele observed in L5b functions similarly to either the ancestral SO1551 allele or L976_03566, but it would be consistent with my motility data given what others have reported for how PDEs and DGCs affect this phenotype (Wei et al., 2016; Liu eta l., 2010; Chao et al., 2013; Jones et al., 2015).

Figure 18: Amino acid conservation between MR-1 SO1551 (top) and Shew-Z12 L976_03566 (bottom). Numbers on either side, represent a scale to indicate amino acid position. The color coding at each position indicates the degree of conservation. Amino acids with similar biochemistry are bluer, while those with divergent biochemistry are more yellow.

Additionally, a survey of the amino acid sequences encoded by both MR-1 and Shew-Z12 genomes revealed that MR-1 had a slightly higher GGDEF:EAL ratio than Shew-Z12 (MR-1 ratio: 1.7 (51:30); Shew-Z12 ratio: 1.64 (54:33); Appendix Table 3). L5b's putative loss of function allele may thus indicate convergence between L5b and Shew-Z12 in terms of the way each lineage's physiological regulatory networks are organized. One potential way to address this would be to compare the levels of cellular c-di-GMP present in L5b and Shew-Z12 to that of MR-1. Lower levels of c-di-GMP in L5b relative to my ancestral MR-1 strain would be consistent with losing DGC activity while maintaining PDE activity in L5b's evolved version of SO1551, and I would expect this same phenotype to be manifested by Shew-Z12, if having lower levels of c-di-GMP is adaptive.

L3a and L5b have not evolved higher colonization densities:

In my system, there are several potential routes to improved fitness. One obvious way to achieve elevated competitive fitness would be to evolve a higher peak abundance in the gut. If an evolved strain were capable of colonizing a larval gut with greater cell densities than its ancestor, the evolved strain would achieve higher competitive indices. Additionally, evolving to have higher growth rates *in vivo*, upon colonization of the larval gut would imply more efficient use of the resources that are available in the gut. There are also less intuitive ways to for an evolved strain to increase its competitive fitness, as improved host association in my system need not rely solely on *in vivo* dynamics. Instead if one considers the aqueous environment outside the larvae, assuming that competing populations have an equal chance of colonizing a larval zebrafish upon encountering this host, one competitor could gain an upper hand by increasing the frequency with which it encountered a host. Further, given equal chances of encountering a host, a passaged lineage might evolve mechanisms that increase the efficiency with which it is able to bypass host filters to translocate from the external aqueous environment into the gut.

With these adaptive strategies in mind, I first sought to understand whether there were differences in the densities L3a and L5b could achieve in the gut. To assess this, I compared the total colony forming units I observed in competitive assays where the ancestor was competed against itself, as well as against L3a and L5b. If L3a or L5b had evolved to achieve greater densities in the larval guts, I would expect to count more CFUs/gut when either competitor was present compared to the ancestor versus ancestor control. I did not observe any increase in carrying capacity when L3a or L5b were present

(Figure 19), suggesting that the competitive performance of these isolates must stem from other components of fitness.

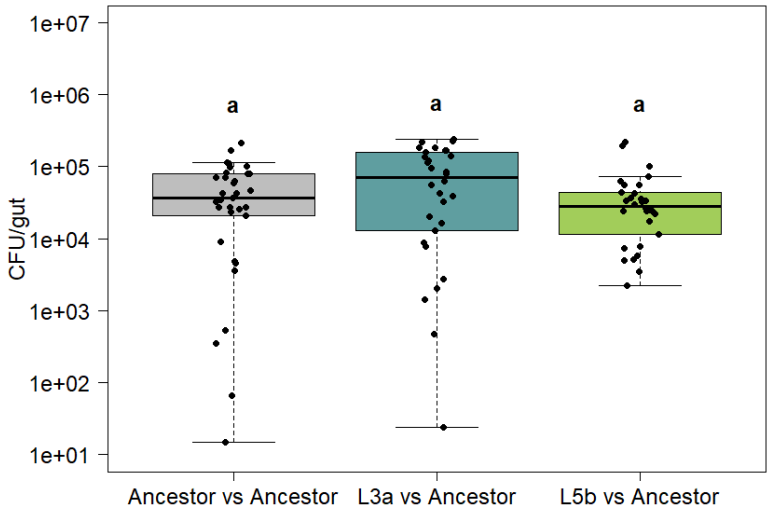

Figure 19: Colonization density achieved in larval guts after 72 hours of colonization under competitive conditions. Dissected guts were dilution plated on TSA and CFUs were counted. Each point represents a single dissected gut. Statistical groupings are indicated by letters above each box for a significance threshold of $p<0.05$. Letters in common between groups indicate the absence of a significant difference in each group's mean.

L5b appears to outcompete the ancestor in vivo:

Despite the fact that neither L3a nor L5b displayed evidence of increased colonization density, I wondered whether L3a and L5b might differ in their growth dynamics *in vivo* when in competition with the ancestor. To assess this, I generated three flasks containing groups of larval zebrafish that had been colonized at a low level (<500 colony forming units (CFUs) per gut) with competition mixtures of either L3a versus the ancestor, or L5b versus the ancestor. I then dissected out the guts of 10 larval zebrafish from three separate flasks per competition and plated the contents of each gut on TSA plates to establish an initial mean competition ratio (CFU/gut_evolved:CFU/gut_ancestor) across the sampled larvae. Each flask contained

greater than 20 larvae, so at least 10 additional larvae remained after the first 10 were sacrificed. Presumably, these fish would have also started out with similar competition ratios to the 10 larvae whose guts I dissected. I then allowed 10 hours to pass (during which time I repeatedly exchanged the EM), so that gut populations could grow and compete within the remaining 10 larvae in each flask. At the end of this period, I again dissected out and plated the guts of 10 additional larvae to quantify a final mean competition ratio per flask. Finally, I calculated a relative fitness metric that was based the ratio of the per capita growth rates of each competing strain (see methods). When I compare the performance of L3a and L5b using this assay, I see that while L3a shows no ability to outcompete the ancestor ($p = 0.612$, $W_{relative} = 0.014 \pm 0.039$, $mean(\log_{10}(W_{relative})) \pm SD(\log_{10}(W_{relative}))$; Figure 20), L5b shows a modest ability to outcompete the ancestor ($p = 0.055$, $W_{relative} = 0.022 \pm 0.009$, $mean(\log_{10}(W_{relative})) \pm SD(\log_{10}(W_{relative}))$; Figure 20). In this assay, if two competitors competed equally with each other, their $\log_{10}(W_{relative})$ value should be exactly zero. Therefore, to assess significance, I t-tested each competition against a μ value of zero.

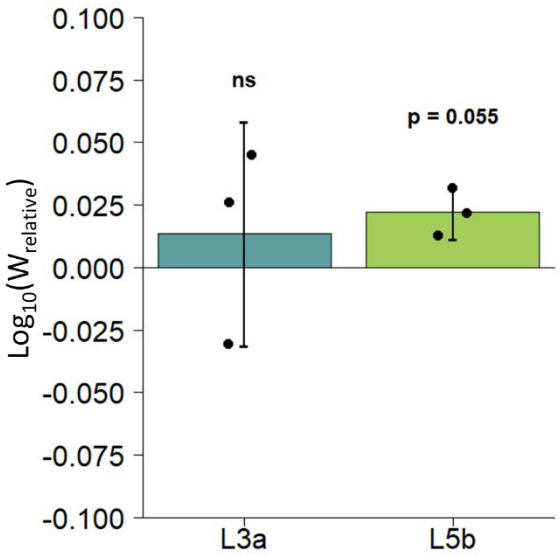

Figure 20: Competitive fitness of L3a and L5b *in vivo*. Competitive *in vivo* growth rates were calculated as described under the *in vivo* growth assay section in the methods (ratio of Malthusian parameters). Each point represents the mean $W_{relative}$ for a single competition. The $\log_{10}(W_{relative})$ for each group of points was t-tested against a mu value of zero, my expected value if there were no competitive advantage.

L5b's growth in vivo growth advantage cannot fully explain its fitness:

While my results suggest that L5b might have a slight growth rate advantage over the ancestral strain, it did not appear there was enough of an advantage to explain the full fitness gain of L5b. To convince myself of this, I generated a discrete Lotka-Volterra model of competition that involved stochastic gut population collapses and then plotted the competitive indices that would result from 72-hour competitions in the gut (Figure 21A). Wiles et al. 2016, had previously shown that simplified bacterial communities periodically undergo stochastic population collapses induced by larval zebrafish peristaltic-like gut motility. Assuming the same dynamics hold true for MR-1 populations in the larval gut, population collapses should provide a strain with a growth advantage the opportunity to amplify its representation in the gut each time a collapse occurs. A representative example of what this might look like can be seen in Figure 21B. This

representation was parameterized assuming that evolved and ancestral strains against each other equally, and it relied on carrying capacity data from my whole system competitions, as well as growth rates inferred from my L5b versus ancestor *in vivo* competitions. To isolate the effects that competitive dynamics solely in the gut would have on L5b's competitive index, I set founding populations of L5b and the ancestor equal to one another, and the founding populations were of a similar magnitude to what I observed in my *in vivo* competition assays (~200 CFU/gut). I simulated the effects of 1000 iterations of this model on L5b's competitive index and found that competitive dynamics in the gut were unlikely to explain L5b's full competitive advantage (*in vivo* simulated CI = 1.5 $\pm$ 0.1, whole system empirical CI = 161.4 $\pm$ 141.9 (Figure 14); mean $\pm$ SD).

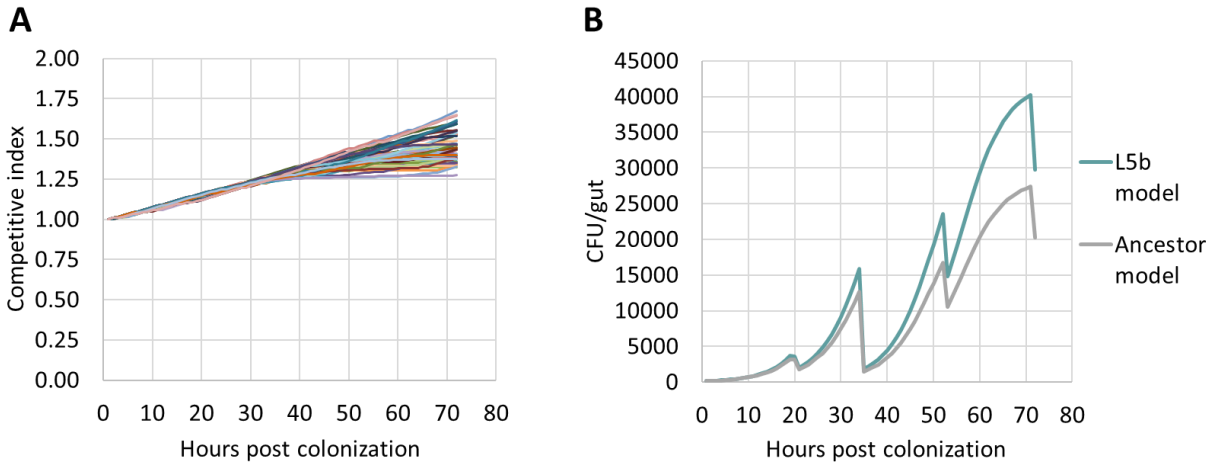

Figure 21: Model of competitive dynamics between L5b and the ancestor *in vivo* over 72 hours. A) Competitive indices resulting from a Lotka-Volterra model of competition that included stochastic *in vivo* population collapses. B) A representative iteration of that model showing the population dynamics for two competing strains in the gut.

L3a and L5b are more represented in the water column, and immigrate more efficiently:

Since *in vivo* competition dynamics alone likely could not explain fitness differences between L3a and L5b, I next investigated how these strains competed against the ancestor outside the larval gut. I accomplished this by competing each evolved isolate against the ancestor in LCM. In these experiments, I plated samples of the LCM at the time of inoculation and at the end of each experiment to establish evolved strain to ancestral strain ratios based on the CFUs for each strain. CIs were then calculated by dividing the ending ratio by the starting ratio. Via this assay, I observed that both L3a and L5b were able to outcompete the ancestor to a similar degree, implying that both evolved strains were able to achieve higher cell densities in the water column (Figure 22A). Having elevated cell densities in the planktonic portion of the LCM (as opposed to being associated with the flask surface) should increase the relative encounter rate between evolved cells and larvae since the larvae are also planktonic.

While having an increased encounter rate with this larval host provides an intuitive path towards improved fitness, the efficiency with which a strain is able to colonize the gut after encountering a larval host could also play a role. I assessed this by adding BF larvae to flasks that contained LCM-cultured competition mixtures, incubating larvae with these competition mixtures for 40-60 minutes so each larva could be colonized, and then dissecting out and plating the contents of 10 colonized larval guts. I then calculated immigration indices by dividing the plated-gut CFUs for each competitor by the competition ratio at the time larvae were added and adjusted for any growth that may have occurred, *in vivo*, during the colonization period. By accounting for the competition ratio larvae were first exposed to, this immigration assay demonstrated that

L3a and L5b both share the ability to immigrate into the larval gut more efficiently than the ancestor on a per capita basis, and that L5b could perform this task significantly better than L3a (Figure 22B). Together, the results of my LCM and immigration competition assays demonstrate that both L3a and L5b are able to outcompete their unpassaged ancestor by increasing their representation in the water column (thereby increasing their host encounter rate), and by more efficiently translocating from the aqueous external environment into the larval gut.

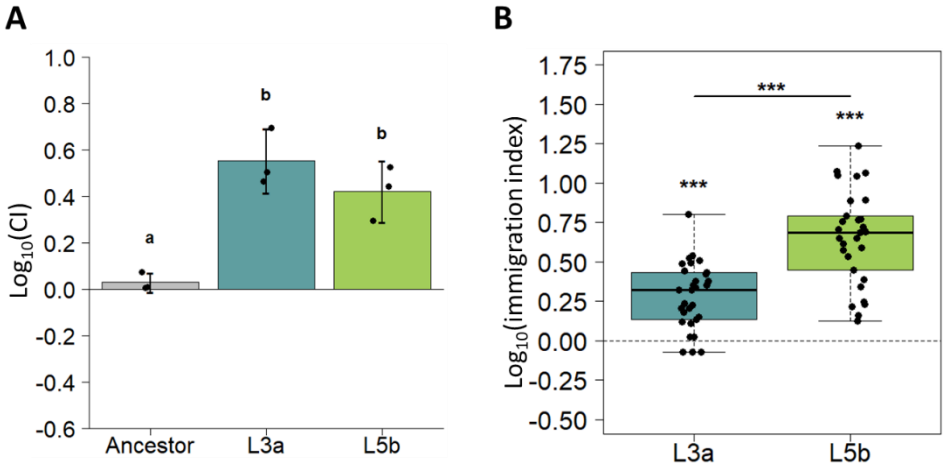

Figure 22: L3a and L5b competitive fitness in larvae-conditioned media and relative per capita immigration compared to ancestor. A) Competitive ability of L3a and L5b against the MR-1 ancestral reference strain in LCM. Each point represents the CI measured for a single LCM competition flask. An ancestral competition against itself is shown as a control to represent the absence of a competitive advantage. Statistical groupings are indicated by letters above each box for a significance threshold of $p<0.05$. Letters in common between groups indicate the absence of a significant difference in each group's mean. B) Boxes show the per capita immigration efficiency of the indicated strain relative to the ancestor. Each point represents a single dissected and plated larval gut. The $\log_{10}$(immigration index) for each group of points was t-tested against a mu value of zero, my expected value if there were no competitive advantage, and Bonferroni corrected for multiple comparisons.

Discussion:

Experimental evolution is a powerful approach to understand how bacteria evolve under a range of different conditions, and most studies that have experimentally evolved bacteria have observed parallelism at the level of fitness (Van den Bergh et al., 2018). Drilling down into lower levels, studies have also observed phenotypic convergence between replicate evolving populations (Chou & Marx, 2012; Tenaillon et al., 2012), and in some cases, even genotypic parallelism has been observed (Bull et al., 1997; Wichman et al., 1999, Good et al., 2017). Here, I conducted an experimental evolution experiment aimed at adapting replicate populations of a free-living bacterial species to a model vertebrate host in order to understand how free-living bacteria initiate transitions towards host association. In both my prior work and in this study, consistent with expectations from past experimental evolution studies (Van den Bergh et al., 2018), I observed that separately evolving replicate populations of MR-1 produced adaptive isolates after serial passage through larval zebrafish. Despite having previously observed high levels of parallel evolution at the level of the mshOP, in this study I aimed to determine whether evolved isolates with divergent adaptive genotypes had convergent adaptive phenotypes.

To assess this I, examined the phenotypes of an evolved isolate containing a mutation in the mshOP of MR-1 (L3a, MshL-T300P) and an evolved isolate that contained a mutation in a putative PDE-DGC domain-containing gene (L5b, SO1551). At the level of fitness, assessed via competition against the ancestor, I observed that both L3a and L5b had improved fitness after serial passage. In considering several components of fitness, I observed that L3a and L5b had a large degree of overlap. Both isolates were capable of outcompeting the ancestor for representation in the water column

where larvae reside, and both isolates had higher per capita immigration rates than the ancestor. I did also observe some differences between these two isolates in that L5b may have a slight growth rate advantage *in vivo* compared to the ancestor, where L3a had none, and that L5b outperformed L3a in terms of its relative immigration ability. These two potential advantages might explain why L5b had higher overall fitness than L3a (Figure 14).

With respect to two behavioral phenotypes I assessed, it was apparent that L3a and L5b are divergent in their biofilm forming capacity under the conditions I tested. Here, and in Chapter II, I demonstrated that L3a's mshOP loss of function mutation is associated with a reduced biofilm forming capacity. This comports well with the growing body of literature suggesting that in mshOP containing bacteria, msh pilus expression is critical to biofilm formation (Saville et al., 2010; Jones et al., 2015; Qin et al., 2014). Interestingly, previous work in *V. cholerae* has shown that an excess of c-di-GMP can increase the expression of the msh pilus through an interaction between c-di-GMP and MshE (the extension ATPase of the msh pilus; Jones et al., 2015), and in MR-1, Thormann and colleagues (2006) demonstrated that reduced activity in a DGC gene (causing a reduction of c-di-GMP) can decrease biofilm formation. Because the deletion mutation I observed L5b's putative DGC domain-containing gene (SO1551) is most likely a loss of function mutation, and yet I do not see a reduced biofilm phenotype in this isolate, I propose SO1551 is not involved in the expression of the msh pilus. However, determining this definitively will require further study.

Puzzlingly, despite the fact that motility often trades off with biofilm formation (Simm et al. 2004, Van Ditmarsch et al., 2013), both L3a and L5b showed increased

motility compared to the ancestor. It is worth noting however, that because the cell densities I employ to foster detectable biofilms are several orders of magnitude higher than what would be experienced by MR-1 cells during a competition assay, it is possible that the biofilms I observe are not representative of the behavior evolved MR-1 isolates manifest when competing to colonize the larval gut. Still, the fact that I see a difference in biofilm formation, as a result of two distinct types of mutations, implies that these mutations at least have the potential to result in unique adaptive trajectories.

Conversely, it is compelling that I observed selection for enhanced motility via two separate genetic pathways in my system, and that this phenotype is also observed in a closely-related *Shewanella* zebrafish isolate (Shew-Z12). When viewed in conjunction with the work others have done (Robinson et al., 2018; Raina et al., 2019; Stephens et al., 2015; Van der Marel et al., 2008), this further establishes that motility is an important factor for host colonization. Although it is not clear exactly how motility improves host colonization in my system, one possibility is that enhanced motility might increase the pace with which chemotactic MR-1 cells traverse host-produced gradients, thereby increasing their encounter rate. Presumably, given that zebrafish larvae are much larger and can move much faster than their potential MR-1 colonists, when larvae are actively swimming, it would alter flow dynamics in the immediate vicinity of the host in ways that would overwhelm the ability of chemotactic responses to influence host encounters (Raina et al., 2019). However, anecdotally zebrafish larvae spend a large portion of their time motionless when they are cohoused in the culture flasks I use in my experiments. During these times of inactivity, chemotactic responses could dramatically increase colonization rates of motile MR-1 cells. My results are consistent with the findings

elsewhere that have shown motility to be an important factor in maintaining host-microbe symbioses (Robinson et al., 2018; Raina et al., 2019; Stephens et al., 2015; Van der Marel et al., 2008).

While I initially speculated that clonal interference might restrict adaptive trajectories through strong selection on a single genomic region, the evidence I present here suggests that there are multiple ways a free-living bacterium could alter its genome to increase its likelihood of becoming host associated. In this study, my experimental evolution approach selected two distinct genotypes with high degrees of phenotypic convergence in a motility phenotype that was associated with increased fitness. However, based on the differences I observed in biofilm formation and some of the components of each genotype's fitness, I suggest that each of the evolved isolates I tested is likely achieving their adaptive gains via different physiological mechanisms. Based on my findings, I propose that fruitful areas for future research would be to address how the mutations I observed impact the expression of cellular components within MR-1 and determine how those cellular components impact the ability of MR-1 populations to exploit specific external and internal niches presented by my host system.

CHAPTER IV:

CONCLUSION

Bacteria were the first known inhabitants of Earth, and since their inception on this planet, they have evolved an incredible diversity of traits. Consequently, they have had a massive impact on the way nutrients cycle within the Earth's biosphere, and all life, in some way, owes its existence to bacteria and the processes they mediate. However, bacteria are not simply bioreactors, reducible solely to their metabolic capacity, they have also evolved a plethora of interesting behaviors which allow them to interact with and occupy all of Earth's unique habitats (Ferenci, 2016; Alivisatos et al., 2015; Thompson et al., 2017). While animals arose out of a world teeming with bacterial life (McFall-Ngai et al., 2013), given how long animals and bacteria have been diverging, for most bacterial species, animals now represent novel, nutrient-rich substrates on and within which bacteria can propagate their future generations. An excellent example of this stems from the ubiquity of animal-bacteria symbioses that I observe around the globe. Exposure to bacteria, and the ability cultivate bacterial communities within their midst, are absolutely critical to animal fitness (Flint et al., 2012; Sommer & Bäckhed, 2013; Dominguez-Bello et al., 2019), yet it is still unknown how microbial communities assemble in animal digestive tracts, and what constrains the evolution of traits that enable bacteria to take part in that assembly. Are there singular traits, that if evolved, dramatically enhance a microbe's ability to colonize a host? Or, does host colonization depend on many traits working in concert? What is the topology of the genetic networks that support these traits? And further, what potentiates the organization of these genetic networks? To address these questions, I experimentally adapted replicate populations of a non-host-

associated bacterium to a model vertebrate host. Specifically, my approach involved the serial passage of *S. oneidensis* (MR-1) populations through the digestive tracts of larval zebrafish, and I chose this approach to select on genotypes that would improve the ability of experimentally evolved populations to colonize larval guts compared to their unpassaged ancestor.

Experimental evolution has a well-documented history of providing insights into the dynamics of evolution under various conditions, and most experimental evolution studies produce organisms that have improved fitness in the environment being considered (Van den Bergh et al., 2018). With the advent of genome sequencing technologies, experimental evolution can be used as a tool to select adaptive genotypes, and then genomic sequencing can be used to identify genetic elements that may be responsible for increases in fitness (Long et al., 2015). Further work in the laboratory can then be conducted to map how specific adaptive genotypes alter behaviors that are associated with improved fitness (Plucain et al., 2014). In this way, the results from experimental-evolution-based studies can help identify the most potent selective pressures imposed by a given environment. In my system I wanted to understand how the selective pressures presented by a host system might facilitate a free-living bacterium's transition to a host associated-existence.

My experimental system consisted of groups of BF larval zebrafish, cohoused in a culture flask at an approximate density of one larva per mL. To construct this environment, I derived zebrafish embryos BF, deposited them in to flasks containing sterile EM, and then allowed those embryos to develop into larvae over the course of 96 hours. During this time, embryos would hatch from their chorions, and begin seeding

their flasks with nutrients. These host-derived nutrients would then be the only nutrients available to sustain MR-1 populations once they were added to these host-containing flasks. In thinking about the full breadth of niche space available in this experimental system, it initially seemed that there could have been an array of strategies free-living bacteria could pursue to improve their ability to colonize larval hosts. For instance, from passage to passage, MR-1 populations would need to grow in the aqueous, extra-host environment to a density where they would frequently encounter and then colonize zebrafish larvae. Once inside larvae, MR-1 cells would need to have traits that enabled their viability in the gut long enough to make it to the next passage. In the next passage, cells would necessitate traits that facilitated their liberation from dissected gut tissue, so they could colonize subsequent hosts. Given that I assessed fitness primarily through competition against an unpassaged ancestral reference strain, MR-1 lineages could have adapted to my system in a number of ways including faster internal or external growth rates, changes in biogeography both inside and outside hosts, altered metabolism, active detection of and navigation towards larval hosts, improved interaction with host tissue, or some combination thereof. Refinement of traits related to any of these aspects of my system could have resulted in the ability to outcompete the ancestor. Thus, it was impossible to know a priori which facet of my system would provide the easiest route towards increased fitness, or whether there were multiple ways to achieve adaptive gains of similar value. Further, given that MR-1 had no history with a vertebrate host, I was unsure how its naivety with zebrafish might bias its evolutionary trajectory.

I found that replicate passaged MR-1 populations exhibited a striking degree of parallelism. At the genomic level, 20 of the 24 isolates I sequenced (four isolates per

replicate population) had mutations in a pilus-encoding operon. The remaining four isolates had mutations in genes that presumably modulated the levels of ci-di-GMP in MR-1 cells. ci-di-GMP is a molecule that many bacteria use to toggle between physiological states, resulting in distinct bacterial behaviors (Römling et al., 2013). In two of the isolates I characterized in depth, both types of mutations resulted in an ability for evolved isolates to outcompete their ancestor in terms of their representation outside larval hosts, as well as in the relative efficiency with which they migrated from the aqueous portion of the environment into the larval gut. Additionally, both types of mutations yielded bacterial strains that were more motile than their ancestor. Although none of these attributes seemed to provide an advantage to evolved lineages once they colonized zebrafish larvae, together they allowed evolved trains to achieve approximately 10-100-fold advantages in terms of their representation in a larval gut when competing against their ancestor.

Despite the parallelism I observed, I did also find subtle phenotypic differences between the two types of mutants I characterized. Namely, L5b had higher fitness overall, likely stemming from an increased immigration efficiency and potentially higher *in vivo* growth rates, and L3a and L5b had different biofilm phenotypes under the conditions I tested. Given these findings, it is interesting to speculate how these differences might structure the adaptive trajectories of evolving lineages if I was to continue my serial passage protocol. Evolutionary theory suggests that the further an organism is from a fitness optimum, the more likely any mutation is to be beneficial (Schoustra et al., 2009; Lenski et al., 2015). Consistent with this, in evolving lineages that are pursuing a fitness optimum, mutations that impart the largest gains towards such an optimum will have the

greatest probability of fixation (Gerrish & Lenski, 1998). In my system, the fact that I consistently saw improvements in the same components of fitness, suggests that the traits that contributed to these improvements were far from their optimal value. In future passages, if the traits that have already improved in my experiment are still far from their optimal value, further optimization could take place resulting in greater gains in the components of fitness in which I have already observed improvements. However, if my selection scheme has resulted in traits that are near their optimal value for certain components of fitness, mutations that optimize the value of traits related to other components of fitness may start to increase in frequency, altering the behavior of the evolving lineage. Additionally, as organisms change, the accessibility of different optima can also change, creating new adaptive pathways, or restricting the likelihood of pursuing a previously likely adaptive pathway. In considering these evolutionary dynamics, given the wide breadth of niche space available in my system, it is possible that additional passages could result in phenotypes that are advantageous to life inside the gut, rather than traits that augment transmission into hosts. Conversely, if there is still a lot of room to improve transmission traits, I might find that motility, immigration, and external growth could be further refined. Moreover, it is possible that although the two types of mutations I examined in my studies had overlapping phenotypes, they could potentiate different adaptive pathways resulting in substantially different phenotypes after continued evolution under my selection scheme. Ultimately, where adaptive isolates lie in relation to their respective fitness optima, should weigh heavily on their evolutionary fate.

Given that bacteria have a wide diversity of traits, and that encounters with some bacteria can fundamentally alter animal well-being, it is important to understand how

bacteria evolve the ability to colonize hosts, thereby providing them with an opportunity to impact host fitness. In my work, I demonstrated that free-living bacteria were able to improve their association with a model vertebrate primarily by enhancing their fitness in the host-constructed external environment. In light of this, I advocate that future microbiome studies consider how bacteria can adapt to those aspects of a host's environment with which a host frequently interacts. If my results are generalizable, I expect that host-associated bacteria likely use environmental components as adaptive bridges which facilitate host colonization, and that bacteria that are well adapted to a host's local environment will evolve novel host symbioses with the greater frequency.

IMG Gene ID	Locus Tag	Gene Symbol	Product Name	DNA Seq Length	Genome
650804462	B565_r001	16S	(rRNA)	1550	Aeromonas veronii B565
2510560287	ACDC_00040000	16S	(rRNA)	1427	Shewanella algae ACDC
2587183130	JCM21037DRAFT_05211	16S	(rRNA)	1427	Shewanella algae JCM 21037
2518739196	Sama_R0001	16S	(rRNA)	1537	Shewanella amazonensis SB2B
648727208	Sbal175DRAFT_R0047	16S	(rRNA)	1533	Shewanella baltica BA175
651216762	Sbal117_R0001	16S	(rRNA)	1533	Shewanella baltica OS117
640723547	Sbal_R0001	16S	(rRNA)	1537	Shewanella baltica OS155
648706592	Sbal183DRAFT_R0088	16S	(rRNA)	1533	Shewanella baltica OS183
641286429	Sbal195_R0007	16S	(rRNA)	1537	Shewanella baltica OS195
2507441981	Sbal625_0042	16S	(rRNA)	1427	Shewanella baltica OS625
2568525539	L876DRAFT_10000	16S	(rRNA)	1421	Shewanella colwelliana ATCC 39565
2524203608	Sden_R0001	16S	(rRNA)	1535	Shewanella denitrificans OS217
2568525540	L884DRAFT_10000	16S	(rRNA)	1421	Shewanella fidelis ATCC BAA-318
641550787	Shal_R0001	16S	(rRNA)	1535	Shewanella halifaxensis HAW-EB4
2566291146	JCM14758DRAFT_05478	16S	(rRNA)	1427	Shewanella haliotis JCM 14758
2521615716	PV4_R0085	16S	(rRNA)	1537	Shewanella loihica PV-4
641234550	Spea_R0001	16S	(rRNA)	1535	Shewanella pealeana ANG-SQ1, ATCC 700345
643423502	swp_rRNA11	16S	(rRNA)	1529	Shewanella piezotolerans WP3
2510452616	Sput200_R0130	16S	(rRNA)	1531	Shewanella putrefaciens 200 (Missing data)
2524134011	Sputcn32_R0001	16S	(rRNA)	1535	Shewanella putrefaciens CN-32
2587191865	JCM20190DRAFT_04395	16S	(rRNA)	1423	Shewanella putrefaciens JCM 20190
640930869	Ssed_R0001	16S	(rRNA)	1535	Shewanella sediminis HAW-EB3
640720070	Shewana3_R0001	16S	(rRNA)	1535	Shewanella sp. ANA-3
640717202	Shewmr4_R0001	16S	(rRNA)	1535	Shewanella sp. MR-4
640717343	Shewmr7_R0001	16S	(rRNA)	1535	Shewanella sp. MR-7
2551868620	SHEWPOL2DRAFT_04550	16S	(rRNA)	1425	Shewanella sp. POL2
640721766	Sputw3181_R0001	16S	(rRNA)	1535	Shewanella sp. W3-18-1
646719986	SVI_r001	16S	(rRNA)	1545	Shewanella violacea DSS12
641631765	Swoo_R0007	16S	(rRNA)	1535	Shewanella woodyi MS32, ATCC 51908
642782354	VFMJ11_A1267	16S	(rRNA)	1552	Vibrio fischeri MJ11
647180815	VFA_rna081	16S	(rRNA)	1546	Vibrio furnissii sv. II CIP 102972

Table 2: Metadata for 16S genes used to create the phylogenetic tree featured in Figure 2

Function	Function Name	Gene Count	Genome Name
pfam00990	GGDEF - Diguanylate cyclase, GGDEF domain	51	Shewanella oneidensis MR-1
pfam00990	GGDEF - Diguanylate cyclase, GGDEF domain	54	Shewanella sp. ZOR0012
pfam00563	EAL - EAL domain	30	Shewanella oneidensis MR-1
pfam00563	EAL - EAL domain	33	Shewanella sp. ZOR0012

Table 3: Metadata for genes containing GGDEF and EAL domains in *S. oneidensis* MR-1 and *Shewanella* sp. ZOR0012 genomes

REFERENCES CITED

Alivisatos, A. P., Blaser, M. J., Brodie, E. L., Chun, M., Dangl, J. L., Donohue, T. J., ... & Knight, R. (2015). A unified initiative to harness Earth's microbiomes. *Science*, *350*(6260), 507-508

Antoniani, D., Bocci, P., Maciąg, A., Raffaelli, N., & Landini, P. (2010). Monitoring of diguanylate cyclase activity and of cyclic-di-GMP biosynthesis by whole-cell assays suitable for high-throughput screening of biofilm inhibitors. *Applied microbiology and biotechnology*, *85*(4), 1095-1104.

Ariyakumar, D. S., & Nishiguchi, M. K. (2009). Characterization of two host-specific genes, mannose-sensitive hemagglutinin (mshA) and uridyl phosphate dehydrogenase (UDPDH) that are involved in the Vibrio fischeri–Euprymna tasmanica mutualism. *FEMS microbiology letters*, *299*(1), 65-73.

Arredondo-Alonso, S., Top, J., Schürch, A. C., McNally, A., Puranen, S., Pesonen, M., ... & van Schaik, W. (2019). Genomes of a major nosocomial pathogen Enterococcus faecium are shaped by adaptive evolution of the chromosome and plasmidome. *bioRxiv*, 530725.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., ... & Meyer, F. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC genomics*, *9*(1), 75.

Bailey, S. F., & Kassen, R. (2012). Spatial structure of ecological opportunity drives adaptation in a bacterium. *The American Naturalist*, *180*(2), 270-283.

Bakke, I., Coward, E., Andersen, T., & Vadstein, O. (2015). Selection in the host structures themicrobiota associated with developing cod larvae (G adus morhua). *Environmental microbiology*, *17*(10), 3914-3924.

Barrett, R. D., & Bell, G. (2006). The dynamics of diversification in evolving Pseudomonas populations. *Evolution*, *60*(3), 484-490.

Barrick, J. E., & Lenski, R. E. (2013). Genome dynamics during experimental evolution. *NatureReviews Genetics*, *14*(12), 827.

Barroso-Batista, J., Sousa, A., Lourenço, M., Bergman, M. L., Sobral, D., Demengeot, J., ... & Gordo, I. (2014). The first steps of adaptation of Escherichia coli to the gut are dominated by soft sweeps. *PLoS genetics*, *10*(3), e1004182.

Bates, J. M., Mittge, E., Kuhlman, J., Baden, K. N., Cheesman, S. E., & Guillemin, K. (2006). Distinct signals from the microbiota promote different aspects of zebrafish gut differentiation. *Developmental biology*, *297*(2), 374-386.

Bercik, P., Denou, E., Collins, J., Jackson, W., Lu, J., Jury, J., ... & Verdu, E. F. (2011). The intestinal microbiota affect central levels of brain-derived neurotropic factor and behavior in mice. *Gastroenterology*, *141*(2), 599-609.

Blount, Z. D., Borland, C. Z., & Lenski, R. E. (2008). Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. *Proceedings of the National Academy of Sciences*, *105*(23), 7899-7906.

Bordenstein, S. R., & Theis, K. R. (2015). Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoS biology*, *13*(8), e1002226.

Bouhenni, R. A., Vora, G. J., Biffinger, J. C., Shirodkar, S., Brockman, K., Ray, R., ... & Fitzgerald, L. A. (2010). The role of Shewanella oneidensis MR-1 outer surface structures in extracellular electron transfer. *Electroanalysis: An International Journal Devoted to Fundamental and Practical Aspects of Electroanalysis*, *22*(7-8), 856-864.

Bouskra, D., Brézillon, C., Bérard, M., Werts, C., Varona, R., Boneca, I. G., & Eberl, G. (2008). Lymphoid tissue genesis induced by commensals through NOD1 regulates intestinal homeostasis. *Nature*, *456*(7221), 507.

Brennan, C. A., Mandel, M. J., Gyllborg, M. C., Thomasgard, K. A., & Ruby, E. G. (2013). Genetic determinants of swimming motility in the squid light-organ symbiont Vibrio fischeri. *Microbiology open*, *2*(4), 576-594.

Bright, M., & Bulgheresi, S. (2010). A complex journey: transmission of microbial symbionts. *Nature Reviews Microbiology*, *8*(3), 218.

Broderick, N. A., Buchon, N., & Lemaitre, B. (2014). Microbiota-induced changes in Drosophila melanogaster host gene expression and gut morphology. *MBio*, *5*(3), e01117-14.

Buffie, C. G., Bucci, V., Stein, R. R., McKenney, P. T., Ling, L., Gobourne, A., ... & Littmann, E. (2015). Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. *Nature*, *517*(7533), 205.

Bull, J. J., Badgett, M. R., Wichman, H. A., Huelsenbeck, J. P., Hillis, D. M., Gulati, A., ... & Molineux, I. J. (1997). Exceptional convergent evolution in a virus. *Genetics*, *147*(4), 1497-1507.

Burns, A. R., & Guillemin, K. (2017). The scales of the zebrafish: host–microbiota interactions from proteins to populations. *Current opinion in microbiology*, *38*, 137-141.

Cash, H. L., Whitham, C. V., Behrendt, C. L., & Hooper, L. V. (2006). Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science*, *313*(5790), 1126-1130.

Chao, L., Rakshe, S., Leff, M., & Spormann, A. M. (2013). PdeB, a cyclic Di-GMP-specific phosphodiesterase that regulates Shewanella oneidensis MR-1 motility and biofilm formation. *Journal of bacteriology*, *195*(17), 3827-3833.

Chen, I. M. A., Chu, K., Palaniappan, K., Pillay, M., Ratner, A., Huang, J., ... & Smirnova, T. (2018). IMG/M v. 5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic acids research*, *47*(D1), D666-D677.

Chiotti, K. E., Kvitek, D. J., Schmidt, K. H., Koniges, G., Schwartz, K., Donckels, E. A., ... & Sherlock, G. (2014). The valley-of-death: reciprocal sign epistasis constrains adaptive trajectories in a constant, nutrient limiting environment. *Genomics*, *104*(6), 431-437.

Chou, H. H., & Marx, C. J. (2012). Optimization of gene expression through divergent mutational paths. *Cell Reports*, *1*(2), 133-140.

Cira, N. J., Pearce, M. T., & Quake, S. R. (2018). Neutral and selective dynamics in a synthetic microbial community. *Proceedings of the National Academy of Sciences*, *115*(42), E9842-E9848.

De Palma, G., Lynch, M. D., Lu, J., Dang, V. T., Deng, Y., Jury, J., ... & Pinto-Sanchez, M. I. (2017). Transplantation of fecal microbiota from patients with irritable bowel syndrome alters gut function and behavior in recipient mice. *Science translational medicine*, *9*(379), eaaf6397.

Deutschbauer, A., Price, M. N., Wetmore, K. M., Shao, W., Baumohl, J. K., Xu, Z., ... & Arkin, A. P. (2011). Evidence-based annotation of gene function in Shewanella oneidensis MR-1 using genome-wide fitness profiling across 121 conditions. *PLoS genetics*, *7*(11), e1002385.

de Visser, J. A. G., & Rozen, D. E. (2006). Clonal interference and the periodic selection of new beneficial mutations in Escherichia coli. *Genetics*, *172*(4), 2093-2100.

Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., & Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proceedings of the National Academy of Sciences*, *107*(26), 11971-11975.

Dominguez-Bello, M. G., Godoy-Vitorino, F., Knight, R., & Blaser, M. J. (2019). Role of the microbiome in human development. *Gut*, gutjnl-2018.

Donaldson, G. P., Lee, S. M., & Mazmanian, S. K. (2016). Gut biogeography of the bacterial microbiota. *Nature Reviews Microbiology*, *14*(1), 20.

Duncan, S. H., Louis, P., Thomson, J. M., & Flint, H. J. (2009). The role of pH in determining the species composition of the human colonic microbiota. *Environmental microbiology*, *11*(8), 2112-2122.

Fabich, A. J., Leatham, M. P., Grissom, J. E., Wiley, G., Lai, H., Najar, F., ... & Conway, T. (2011). Genotype and phenotypes of an intestine-adapted Escherichia coli K-12 mutant selected by animal passage for superior colonization. *Infection and immunity*, *79*(6), 2430-2439.

Ferenci, T. (2016). Trade-off mechanisms shaping the diversity of bacteria. Trends in microbiology, 24(3), 209-223.

Flint, H. J., Scott, K. P., Louis, P., & Duncan, S. H. (2012). The role of the gut microbiota in nutrition and health. *Nature reviews Gastroenterology & hepatology*, *9*(10), 577.

Flores, M. V., Crawford, K. C., Pullin, L. M., Hall, C. J., Crosier, K. E., & Crosier, P. S. (2010). Dual oxidase in the intestinal epithelium of zebrafish larvae has anti-bacterial properties. *Biochemical and biophysical research communications*, *400*(1), 164-168.

Flynn, K. M., Dowell, G., Johnson, T. M., Koestler, B. J., Waters, C. M., & Cooper, V. S. (2016). Evolution of ecological diversity in biofilms of Pseudomonas aeruginosa by altered cyclic diguanylate signaling. *Journal of bacteriology*, *198*(19), 2608-2618.

Fogle, C. A., Nagle, J. L., & Desai, M. M. (2008). Clonal interference, multiple mutations and adaptation in large asexual populations. *Genetics*, *180*(4), 2163-2173.

Fong, S. S., Joyce, A. R., & Palsson, B. Ø. (2005). Parallel adaptive evolution cultures of Escherichia coli lead to convergent growth phenotypes with different gene expression states. *Genome research*, *15*(10), 1365-1372.

Frenkel, E. M., McDonald, M. J., Van Dyken, J. D., Kosheleva, K., Lang, G. I., & Desai, M. M. (2015). Crowded growth leads to the spontaneous evolution of semistable coexistence in laboratory yeast populations. *Proceedings of the National Academy of Sciences*, *112*(36), 11306-11311.

Frese, S. A., Benson, A. K., Tannock, G. W., Loach, D. M., Kim, J., Zhang, M., ... & MacKenzie, D. A. (2011). The evolution of host specialization in the vertebrate gut symbiont Lactobacillus reuteri. *PLoS genetics*, *7*(2), e1001314.

Friedman, E. S., Bittinger, K., Esipova, T. V., Hou, L., Chau, L., Jiang, J., ... & Goulian, M. (2018). Microbes vs. chemistry in the origin of the anaerobic gut lumen. *Proceedings of the National Academy of Sciences*, *115*(16), 4170-4175.

Fukushima, K., Ogawa, H., Takahashi, K., Naito, H., Funayama, Y., Kitayama, T., ... & Sasaki, I. (2003). Non-pathogenic bacteria modulate colonic epithelial gene expression in germ-free mice. *Scandinavian journal of gastroenterology*, *38*(6), 626-634.

Gao, H., Wang, Y., Liu, X., Yan, T., Wu, L., Alm, E., ... & Zhou, J. (2004). Global transcriptome analysis of the heat shock response of Shewanella oneidensis. *Journal of bacteriology*, *186*(22), 7796-7803.

Gerrish, P. J., & Lenski, R. E. (1998). The fate of competing beneficial mutations in an asexual population. *Genetica*, *102*, 127.

Giraud, A., Matic, I., Tenaillon, O., Clara, A., Radman, M., Fons, M., & Taddei, F. (2001). Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. *science*, *291*(5513), 2606-2608.

Giraud, A., Arous, S., De Paepe, M., Gaboriau-Routhiau, V., Bambou, J. C., Rakotobe, S., ... & Cerf-Bensussan, N. (2008). Dissecting the genetic components of adaptation of Escherichia coli to the mouse gut. *PLoS genetics*, *4*(1), e2.

Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E., & Desai, M. M. (2017). The dynamics of molecular evolution over 60,000 generations. *Nature*, *551*(7678), 45.

Goodman, A. L., McNulty, N. P., Zhao, Y., Leip, D., Mitra, R. D., Lozupone, C. A., ... & Gordon, J. I. (2009). Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell host & microbe*, *6*(3), 279-289.

Goodrich, J. K., Davenport, E. R., Beaumont, M., Jackson, M. A., Knight, R., Ober, C., ... & Ley, R. E. (2016). Genetic determinants of the gut microbiome in UK twins. *Cell host & microbe*, *19*(5), 731-743.

Heidelberg, J. F., Paulsen, I. T., Nelson, K. E., Gaidos, E. J., Nelson, W. C., Read, T. D., ... & Clayton, R. A. (2002). Genome sequence of the dissimilatory metal ion–reducing bacterium Shewanella oneidensis. *Nature biotechnology*, *20*(11), 1118.

Heijtz, R. D., Wang, S., Anuar, F., Qian, Y., Björkholm, B., Samuelsson, A., ... & Pettersson, S. (2011). Normal gut microbiota modulates brain development and behavior. *Proceedings of the National Academy of Sciences*, *108*(7), 3047-3052.

Herron, M. D., & Doebeli, M. (2013). Parallel evolutionary dynamics of adaptive diversification in Escherichia coli. *PLoS biology*, *11*(2), e1001490.

Hou, J., Liu, Y., Liu, Y., & Shao, Y. (2012). The MSHA strain of Pseudomonas aeruginosa activated TLR pathway and enhanced HIV-1 DNA vaccine immunoreactivity. *PloS one*, *7*(10), e47724.

Hsiao, A., Liu, Z., Joelsson, A., & Zhu, J. (2006). Vibrio cholerae virulence regulator-coordinated evasion of host immunity. *Proceedings of the National Academy of Sciences*, *103*(39), 14542-14547.

Hsiao, E. Y., McBride, S. W., Hsien, S., Sharon, G., Hyde, E. R., McCue, T., ... & Patterson, P. H. (2013). Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell*, *155*(7), 1451-1463.

Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., ... & Suzuki, Y. (2016). A new view of the tree of life. *Nature microbiology*, *1*(5), 16048.

Islam, K. S., Fukiya, S., Hagio, M., Fujii, N., Ishizuka, S., Ooka, T., ... & Yokota, A. (2011). Bile acid is a host factor that regulates the composition of the cecal microbiota in rats. *Gastroenterology*, *141*(5), 1773-1781.

Jacob, V., Crawford, C., Cohen-Mekelburg, S., Viladomiu, M., Putzel, G. G., Schneider, Y., ... & Ajami, N. J. (2017). Single delivery of high-diversity fecal microbiota preparation by colonoscopy is safe and effective in increasing microbial diversity in active ulcerative colitis. *Inflammatory bowel diseases*, *23*(6), 903-911.

Jenal, U., & Malone, J. (2006). Mechanisms of cyclic-di-GMP signaling in bacteria. *Annu. Rev. Genet.*, *40*, 385-407.

Jiang, Y., Dong, Y., Luo, Q., Li, N., Wu, G., & Gao, H. (2014). Protection from oxidative stress relies mainly on derepression of OxyR-dependent KatB and Dps in Shewanella oneidensis. *Journal of bacteriology*, *196*(2), 445-458.

Jones, B. V., Begley, M., Hill, C., Gahan, C. G., & Marchesi, J. R. (2008). Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proceedings of the national academy of sciences*, *105*(36), 13580-13585.

Jones, C. J., Utada, A., Davis, K. R., Thongsomboon, W., Sanchez, D. Z., Banakar, V., ... & Yildiz, F. H. (2015). C-di-GMP regulates motile to sessile transition by modulating MshA pili biogenesis and near-surface motility behavior in Vibrio cholerae. *PLoS pathogens*, *11*(10), e1005068.

Khan, A. I., Dinh, D. M., Schneider, D., Lenski, R. E., & Cooper, T. F. (2011). Negative epistasis between beneficial mutations in an evolving bacterial population. *Science*, *332*(6034), 1193-1196.

Kirillina, O., Fetherston, J. D., Bobrov, A. G., Abney, J., & Perry, R. D. (2004). HmsP, a putative phosphodiesterase, and HmsT, a putative diguanylate cyclase, control Hms-dependent biofilm formation in Yersinia pestis. *Molecular microbiology*, *54*(1), 75-88.

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome research*, *27*(5), 722-736.

Koropatkin, N. M., Cameron, E. A., & Martens, E. C. (2012). How glycan metabolism shapes the human gut microbiota. *Nature Reviews Microbiology*, *10*(5), 323.

Kvitek, D. J., & Sherlock, G. (2011). Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. *PLoS genetics*, *7*(4), e1002056.

Larsson, E., Tremaroli, V., Lee, Y. S., Koren, O., Nookaew, I., Fricker, A., ... & Bäckhed, F. (2012). Analysis of gut microbial regulation of host gene expression along the length of the gut and regulation of gut microbial ecology through MyD88. *gut*, *61*(8), 1124-1131.

Leatham, M. P., Stevenson, S. J., Gauger, E. J., Krogfelt, K. A., Lins, J. J., Haddock, T. L., ... & Cohen, P. S. (2005). Mouse intestine selects nonmotile flhDC mutants of Escherichia coli MG1655 with increased colonizing ability and better utilization of carbon sources. *Infection and immunity*, *73*(12), 8039-8049.

Lee, S. M., Donaldson, G. P., Mikulski, Z., Boyajian, S., Ley, K., & Mazmanian, S. K. (2013). Bacterial colonization factors control specificity and stability of the gut microbiota. *Nature*, *501*(7467), 426

Lee, J. Y., Han, G. G., Kim, E. B., & Choi, Y. J. (2017). Comparative genomics of Lactobacillus salivarius strains focusing on their host adaptation. Microbiological research, 205, 48-58.

Lenski, R. E., Rose, M. R., Simpson, S. C., & Tadler, S. C. (1991). Long-term experimental evolution in Escherichia coli. I. Adaptation and divergence during 2,000 generations. *The American Naturalist*, *138*(6), 1315-1341.

Lenski, R. E., Wiser, M. J., Ribeck, N., Blount, Z. D., Nahum, J. R., Morris, J. J., ... & Burmeister, A. R. (2015). Sustained fitness gains and variability in fitness trajectories in the long-term evolution experiment with Escherichia coli. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1821), 20152292.

Lenski, R. E. (2017). Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *The ISME journal*, *11*(10), 2181.

Lescat, M., Launay, A., Ghalayini, M., Magnan, M., Glodt, J., Pintard, C., ... & Tenaillon, O. (2017). Using long-term experimental evolution to uncover the patterns and determinants of molecular evolution of an Escherichia coli natural isolate in the streptomycin-treated mouse gut. *Molecular ecology*, *26*(7), 1802-1817.

Ley, R. E., Bäckhed, F., Turnbaugh, P., Lozupone, C. A., Knight, R. D., & Gordon, J. I. (2005). Obesity alters gut microbial ecology. *Proceedings of the National Academy of Sciences*, *102*(31), 11070-11075.

Ley, R. E., Lozupone, C. A., Hamady, M., Knight, R., & Gordon, J. I. (2008). Worlds within worlds: evolution of the vertebrate gut microbiota. *Nature Reviews Microbiology*, *6*(10), 776.

Li, R., Auchtung, J. M., Tiedje, J. M., & Worden, R. M. (2010). Shewanella oneidensis MR-1 chemotaxis in a diffusion gradient chamber. *Environmental science & technology*, *45*(3), 1014-1020.

Licandro-Seraut, H., Scornec, H., Pédron, T., Cavin, J. F., & Sansonetti, P. J. (2014). Functional genomics of Lactobacillus casei establishment in the gut. *Proceedings of the National Academy of Sciences*, *111*(30), E3101-E3109.

List, C., Grutsch, A., Radler, C., Cakar, F., Zingl, F. G., Schild-Prüfert, K., & Schild, S. (2018). Genes Activated by Vibrio cholerae upon Exposure to Caenorhabditis elegans Reveal the Mannose-Sensitive Hemagglutinin To Be Essential for Colonization. *mSphere*, *3*(3), e00238-18.

Liu, Y., Gao, W., Wang, Y., Wu, L., Liu, X., Yan, T., ... & Zhou, J. (2005). Transcriptome analysis of Shewanella oneidensis MR-1 in response to elevated salt conditions. *Journal of bacteriology*, *187*(7), 2501-2507.

Liu, X., Beyhan, S., Lim, B., Linington, R. G., & Yildiz, F. H. (2010). Identification and characterization of a phosphodiesterase that inversely regulates motility and biofilm formation in Vibrio cholerae. *Journal of bacteriology*, *192*(18), 4541-4552.

Long, A., Liti, G., Luptak, A., & Tenaillon, O. (2015). Elucidating the molecular architecture of adaptation via evolve and resequence experiments. *Nature Reviews Genetics*, *16*(10), 567.

Luczynski, P., McVey Neufeld, K. A., Oriach, C. S., Clarke, G., Dinan, T. G., & Cryan, J. F. (2016). Growing up in a bubble: using germ-free animals to assess the influence of the gut microbiota on brain and behavior. *International Journal of Neuropsychopharmacology*, *19*(8).

Mani, G. S., & Clarke, B. C. (1990). Mutational order: a major stochastic process in evolution. *Proceedings of the Royal Society of London. B. Biological Sciences*, *240*(1297), 29-37.

Mazel, F., Davis, K. M., Loudon, A., Kwong, W. K., Groussin, M., & Parfrey, L. W. (2018). Is Host Host Filtering the Main Driver of Phylosymbiosis across the Tree of Life?. *mSystems*, *3*(5), e00097-18.

Mazmanian, S. K., Liu, C. H., Tzianabos, A. O., & Kasper, D. L. (2005). An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell*, *122*(1), 107-118

McFall-Ngai, M., Hadfield, M. G., Bosch, T. C., Carey, H. V., Domazet-Lošo, T., Douglas, A. E., ... & Hentschel, U. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences*, *110*(9), 3229-3236.

McGuinness, W., Kobayashi, S., & DeLeo, F. (2016). Evasion of neutrophil killing by Staphylococcus aureus. *Pathogens*, *5*(1), 32.

McLean, J. S., Pinchuk, G. E., Geydebrekht, O. V., Bilskis, C. L., Zakrajsek, B. A., Hill, E. A., ... & Beliaev, A. S. (2008). Oxygen-dependent autoaggregation in Shewanella oneidensis MR-1. *Environmental microbiology*, *10*(7), 1861-1876.

McLoughlin, K., Schluter, J., Rakoff-Nahoum, S., Smith, A. L., & Foster, K. R. (2016). Host selection of microbiota via differential adhesion. *Cell host & microbe*, *19*(4), 550-559.

Melancon, E., Canny, S. G. D. L. T., Sichel, S., Kelly, M., Wiles, T. J., Rawls, J. F., ... & Guillemin, K. (2017). Best practices for germ-free derivation and gnotobiotic zebrafish husbandry. In *Methods in cell biology* (Vol. 138, pp. 61-100). Academic Press.

Merritt, J. H., Ha, D. G., Cowles, K. N., Lu, W., Morales, D. K., Rabinowitz, J., ... & O'Toole, G. A. (2010). Specific control of Pseudomonas aeruginosa surface-associated behaviors by two c-di-GMP diguanylate cyclases. *MBio*, *1*(4), e00183-10.

Morgan, X. C., Tickle, T. L., Sokol, H., Gevers, D., Devaney, K. L., Ward, D. V., ... & Bousvaros, A. (2012). Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome biology*, *13*(9), R79.

Myers, C. R., & Nealson, K. H. (1988). Bacterial manganese reduction and growth with manganese oxide as the sole electron acceptor. *Science*, *240*(4857), 1319-1321.

Nilsson, A. I., Kugelberg, E., Berg, O. G., & Andersson, D. I. (2004). Experimental adaptation of Salmonella typhimurium to mice. *Genetics*, *168*(3), 1119-1130.

O'Boyle, N., Houeix, B., Kilcoyne, M., Joshi, L., & Boyd, A. (2013). The MSHA pilus of Vibrio parahaemolyticus has lectin functionality and enables TTSS-mediated pathogenicity. *International Journal of Medical Microbiology*, *303*(8), 563-573.

Ong, W. K., Vu, T. T., Lovendahl, K. N., Llull, J. M., Serres, M. H., Romine, M. F., & Reed, J. L. (2014). Comparisons of Shewanella strains based on genome annotations, modeling, and experiments. *BMC systems biology*, *8*(1), 31.

Ottman, N., Reunanen, J., Meijerink, M., Pietilä, T. E., Kainulainen, V., Klievink, J., ... & Satokari, R. (2017). Pili-like proteins of Akkermansia muciniphila modulate host immune responses and gut barrier function. *PLoS One*, *12*(3), e0173004.

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., ... & Vonstein, V. (2013). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic acids research*, *42*(D1), D206-D214.

Parsons, E. S., Stanley, G. J., Pyne, A. L., Hodel, A. W., Nievergelt, A. P., Menny, A., ... & Bubeck, D. (2019). Single-molecule kinetics of pore assembly by the membrane attack complex. *Nature Communications*, *10*(1), 2066.

Parthasarathy, R. (2012). Rapid, accurate particle tracking by calculation of radial symmetry centers. *Nature methods*, *9*(7), 724.

Plucain, J., Hindré, T., Le Gac, M., Tenaillon, O., Cruveiller, S., Médigue, C., ... & Schneider, D. (2014). Epistasis and allele specificity in the emergence of a stable polymorphism in Escherichia coli. *Science*, *343*(6177), 1366-1369.

Qin, Y. X., Yan, Q. P., Mao, X. X., Chen, Z., & Su, Y. Q. (2014). Role of MshQ in MSHA pili biosynthesis and biofilm formation of Aeromonas hydrophila. *Genet. Mol. Res*, *13*(4), 8982-8996.

Quinn, R. A., Comstock, W., Zhang, T., Morton, J. T., da Silva, R., Tran, A., ... & Ackermann, G. (2018). Niche partitioning of a pathogenic microbiome driven by chemical gradients. *Science advances*, *4*(9), eaau1908.

Raina, J. B., Fernandez, V., Lambert, B., Stocker, R., & Seymour, J. R. (2019). The role of microbial motility and chemotaxis in symbiosis. *Nature Reviews Microbiology*, 1.

Rainey, P. B., & Travisano, M. (1998). Adaptive radiation in a heterogeneous environment. *Nature*, *394*(6688), 69.

Rambaut, A., & Drummond, A. J. (2012). FigTree version 1.4. 0.

Rawls, J. F., Samuel, B. S., & Gordon, J. I. (2004). Gnotobiotic zebrafish reveal evolutionarily conserved responses to the gut microbiota. *Proceedings of the National Academy of Sciences*, *101*(13), 4596-4601.

Rawls, J. F., Mahowald, M. A., Ley, R. E., & Gordon, J. I. (2006). Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. *Cell*, *127*(2), 423-433.

Rey, F. E., Faith, J. J., Bain, J., Muehlbauer, M. J., Stevens, R. D., Newgard, C. B., & Gordon, J. I. (2010). Dissecting the in vivo metabolic potential of two human gut acetogens. *Journal of Biological Chemistry*, *285*(29), 22082-22090.

Ribet, D., & Cossart, P. (2015). How bacterial pathogens colonize their hosts and invade deeper tissues. *Microbes and infection*, *17*(3), 173-183.

Robinson, C. D., Klein, H. S., Murphy, K. D., Parthasarathy, R., Guillemin, K., & Bohannan, B. J. (2018). Experimental bacterial adaptation to the zebrafish gut reveals a primary role for immigration. *PLoS biology*, *16*(12), e2006893.

Rodrigues, J. L., Serres, M. H., & Tiedje, J. M. (2011). Large-scale comparative phenotypic and genomic analyses reveal ecological preferences of Shewanella species and identify metabolic pathways conserved at the genus level. *Appl. Environ. Microbiol.*, *77*(15), 5352-5360.

Rolig, A. S., Parthasarathy, R., Burns, A. R., Bohannan, B. J., & Guillemin, K. (2015). Individual members of the microbiota disproportionately modulate host innate immune responses. *Cell host & microbe*, *18*(5), 613-620.

Römling, U., Galperin, M. Y., & Gomelsky, M. (2013). Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol. Mol. Biol. Rev.*, *77*(1), 1-52.

Rothschild, D., Weissbrod, O., Barkan, E., Kurilshikov, A., Korem, T., Zeevi, D., ... & Shilo, S. (2018). Environment dominates over host genetics in shaping human gut microbiota. *Nature*, *555*(7695), 210.

Rozen, D. E., De Visser, J. A. G., & Gerrish, P. J. (2002). Fitness effects of fixed beneficial mutations in microbial populations. *Current biology*, *12*(12), 1040-1045.

Sachs, J. L., Skophammer, R. G., & Regus, J. U. (2011). Evolutionary transitions in bacterial symbiosis. *Proceedings of the National Academy of Sciences*, *108*(Supplement 2), 10800-10807.

Sarkar, S., Hutton, M. L., Vagenas, D., Ruter, R., Schüller, S., Lyras, D., ... & Totsika, M. (2018). Intestinal colonization traits of Pandemic multidrug-resistant Escherichia coli ST131. *The Journal of infectious diseases*, *218*(6), 979-990.

Sass, V., Schneider, T., Wilmes, M., Körner, C., Tossi, A., Novikova, N., ... & Sahl, H. G. (2010). Human β-defensin 3 inhibits cell wall biosynthesis in Staphylococci. *Infection and immunity*, *78*(6), 2793-2800.

Saville, R. M., Dieckmann, N., & Spormann, A. M. (2010). Spatiotemporal activity of the mshA gene system in Shewanella oneidensis MR-1 biofilms. *FEMS microbiology letters*, *308*(1), 76-83.

Schlomann, B. H., Wiles, T. J., Wall, E. S., Guillemin, K., & Parthasarathy, R. (2018). Bacterial cohesion predicts spatial distribution in the larval zebrafish intestine. *Biophysical journal*, *115*(11), 2271-2277

Schluter, J., & Foster, K. R. (2012). The evolution of mutualism in gut microbiota via host epithelial selection. *PLoS biology*, *10*(11), e1001424.

Schmidt, A. J., Ryjenkov, D. A., & Gomelsky, M. (2005). The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. *Journal of bacteriology*, *187*(14), 4774-4781.

Schoustra, S. E., Bataillon, T., Gifford, D. R., & Kassen, R. (2009). The properties of adaptive walks in evolving populations of fungus. *PLoS biology*, *7*(11), e1000250.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068-2069

Segal, A. W. (2005). How neutrophils kill microbes. *Annu. Rev. Immunol.*, *23*, 197-223.

Semova, I., Carten, J. D., Stombaugh, J., Mackey, L. C., Knight, R., Farber, S. A., & Rawls, J. F. (2012). Microbiota regulate intestinal absorption and metabolism of fatty acids in the zebrafish. *Cell host & microbe*, *12*(3), 277-288.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., ... & Thompson, J. D. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*, *7*(1), 539.

Simm, R., Morr, M., Kader, A., Nimtz, M., & Römling, U. (2004). GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Molecular microbiology*, *53*(4), 1123-1134.

Sommer, F., & Bäckhed, F. (2013). The gut microbiota—masters of host development and physiology. *Nature Reviews Microbiology*, *11*(4), 227.

Spencer, C. C., Tyerman, J., Bertrand, M., & Doebeli, M. (2008). Adaptation increases the likelihood of diversification in an experimental bacterial lineage. *Proceedings of the National Academy of Sciences*, *105*(5), 1585-1589.

Stagaman, K., Burns, A. R., Guillemin, K., & Bohannan, B. J. (2017). The role of adaptive immunity as an ecological filter on the gut microbiota in zebrafish. *The ISME journal*, *11*(7), 1630.

Stephens, W. Z., Wiles, T. J., Martinez, E. S., Jemielita, M., Burns, A. R., Parthasarathy, R., ... & Guillemin, K. (2015). Identification of population bottlenecks and colonization factors during assembly of bacterial communities within the zebrafish intestine. *MBio*, *6*(6), e01163-15.

Stephens, W. Z., Burns, A. R., Stagaman, K., Wong, S., Rawls, J. F., Guillemin, K., & Bohannan, B. J. (2016). The composition of the zebrafish intestinal microbial community varies across development. *The ISME journal*, *10*(3), 644.

Suriyampola, P. S., Shelton, D. S., Shukla, R., Roy, T., Bhat, A., & Martins, E. P. (2016). Zebrafish social behavior in the wild. *Zebrafish*, *13*(1), 1-8.

Taormina, M. J., Hay, E. A., & Parthasarathy, R. (2017). Passive and active microrheology of the intestinal fluid of the larval zebrafish. *Biophysical journal*, *113*(4), 957-965.

Tenaillon, O., Rodríguez-Verdugo, A., Gaut, R. L., McDonald, P., Bennett, A. F., Long, A. D., & Gaut, B. S. (2012). The molecular diversity of adaptive convergence. *Science*, *335*(6067), 457-461.

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., ... & Navas-Molina, J. A. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*, *551*(7681), 457.

Thormann, K. M., Duttler, S., Saville, R. M., Hyodo, M., Shukla, S., Hayakawa, Y., & Spormann, A. M. (2006). Control of formation and cellular detachment from Shewanella oneidensis MR-1 biofilms by cyclic di-GMP. *Journal of bacteriology*, *188*(7), 2681-2691.

Traverse, C. C., Mayo-Smith, L. M., Poltak, S. R., & Cooper, V. S. (2013). Tangled bank of experimentally evolved Burkholderia biofilms reflects selection during chronic infections. *Proceedings of the National Academy of Sciences*, *110*(3), E250-E259.

Turnbaugh, P. J., Ley, R. E., Mahowald, M. A., Magrini, V., Mardis, E. R., & Gordon, J. I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *nature*, *444*(7122), 1027.

Turner, P. E., Souza, V., & Lenski, R. E. (1996). Tests of ecological mechanisms promoting the stable coexistence of two bacterial genotypes. *Ecology*, *77*(7), 2119-2129.

Van den Bergh, B., Swings, T., Fauvart, M., & Michiels, J. (2018). Experimental design, population dynamics, and diversity in microbial experimental evolution. *Microbiol. Mol. Biol. Rev.*, *82*(3), e00008-18.

Van der Marel, M., Schroers, V., Neuhaus, H., & Steinhagen, D. (2008). Chemotaxis towards, adhesion to, and growth in carp gut mucus of two Aeromonas hydrophila strains with different pathogenicity for common carp, Cyprinus carpio L. *Journal of fish diseases*, *31*(5), 321-330.

Van Ditmarsch, D., Boyle, K. E., Sakhtah, H., Oyler, J. E., Nadell, C. D., Déziel, É., ... & Xavier, J. B. (2013). Convergent evolution of hyperswarming leads to impaired biofilm formation in pathogenic bacteria. *Cell reports*, *4*(4), 697-708.

Van Nood, E., Vrieze, A., Nieuwdorp, M., Fuentes, S., Zoetendal, E. G., de Vos, W. M., ... & Speelman, P. (2013). Duodenal infusion of donor feces for recurrent Clostridium difficile. *New England Journal of Medicine*, *368*(5), 407-415.

Venkateswaran, K., Moser, D. P., Dollhopf, M. E., Lies, D. P., Saffarini, D. A., MacGregor, B. J., ... & Burghardt, J. (1999). Polyphasic taxonomy of the genus Shewanella and description of Shewanella oneidensis sp. nov. *International Journal of Systematic and Evolutionary Microbiology*, *49*(2), 705-724.

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189-1191.

Wei, C., Jiang, W., Zhao, M., Ling, J., Zeng, X., Deng, J., ... & Sun, W. (2016). A systematic analysis of the role of GGDEF-EAL domain proteins in virulence and motility in Xanthomonas oryzae pv. oryzicola. *Scientific reports*, *6*, 23769.

Weiss, G., & Schaible, U. E. (2015). Macrophage defense mechanisms against intracellular bacteria. *Immunological reviews*, *264*(1), 182-203.

Wichman, H. A., Badgett, M. R., Scott, L. A., Boulianne, C. M., & Bull, J. J. (1999). Different trajectories of parallel evolution during viral adaptation. *Science*, *285*(5426), 422-424.

Wiles, T. J., Jemielita, M., Baker, R. P., Schlomann, B. H., Logan, S. L., Ganz, J., ... & Parthasarathy, R. (2016). Host gut motility promotes competitive exclusion within a model intestinal microbiota. *PLoS biology*, *14*(7), e1002517.

Wiles, T. J., Wall, E. S., Schlomann, B. H., Hay, E. A., Parthasarathy, R., & Guillemin, K. (2018). Modernized tools for streamlined genetic manipulation and comparative study of wild and diverse proteobacterial lineages. *MBio*, *9*(5), e01877-18.

Wilke, C. O. (2004). The speed of adaptation in large asexual populations. *Genetics*, *167*(4), 2045-2053.

Wiser, M. J., Ribeck, N., & Lenski, R. E. (2013). Long-term dynamics of adaptation in asexual populations. *Science*, *342*(6164), 1364-1367.

Wong, S., Stephens, W. Z., Burns, A. R., Stagaman, K., David, L. A., Bohannan, B. J., ... & Rawls, J. F. (2015). Ontogenetic differences in dietary fat influence microbiota assembly in the zebrafish gut. *MBio*, *6*(5), e00687-15.

Wu, W. M., Yang, Y. S., & Peng, L. H. (2014). Microbiota in the stomach: new insights. *Journal of digestive diseases*, *15*(2), 54-61.

Xu, Q., Dziejman, M., & Mekalanos, J. J. (2003). Determination of the transcriptome of Vibrio cholerae during intraintestinal growth and midexponential phase in vitro. *Proceedings of the National Academy of Sciences*, *100*(3), 1286-1291.

Yahara, K., Méric, G., Taylor, A. J., de Vries, S. P., Murray, S., Pascoe, B., ... & Komukai, S. (2017). Genome-wide association of functional traits linked with C ampylobacter jejuni survival from farm to fork. *Environmental microbiology*, *19*(1), 361-380.

Yoon, S. H., Ha, S. M., Lim, J., Kwon, S., & Chun, J. (2017). A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek*, *110*(10), 1281-1286.

Zhao, S., Lieberman, T. D., Poyet, M., Groussin, M., Gibbons, S. M., Xavier, R. J., & Alm, E. J. (2018). Adaptive evolution within the gut microbiome of individual people. *bioRxiv*, 208009.