INVESTIGATING THE ROLE OF IMMUNITY AND OTHER SELECTIVE PRESSURES ON THE ASSEMBLY OF THE GUT MICROBIOTA IN ZEBRAFISH AND HUMANS

by

KEATON DANIEL STAGAMAN

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 2016

DISSERTATION APPROVAL PAGE

Student: Keaton Daniel Stagaman

Title: Investigating the Role of Immunity and Other Selective Pressures on the Assembly of the Gut Microbiota in Zebrafish and Humans

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:

William Cresko	Chairperson
Brendan Bohannan	Advisor
Karen Guillemin	Advisor
John Postlethwaite	Core Member
J. Josh Snodgrass	Institutional Representative

and

Scott L. Pratt	Dean of the Graduate School

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

Degree awarded June 2016

© 2016 Keaton Daniel Stagaman
This work is licensed under a Creative Commons
Attribution-NonCommercial-NoDerivs (United States) License.

DISSERTATION ABSTRACT

Keaton Daniel Stagaman

Doctor of Philosophy

Department of Biology

June 2016

Title: Investigating the Role of Immunity and Other Selective Pressures on the Assembly of the Gut Microbiota in Zebrafish and Humans

Over the past few decades, it has become increasingly apparent that host-associated microbial communities play an integral role in the development, physiology, and health of their host organisms. All hosts have evolved mechanisms to filter the microbial taxa that comprise their resident intestinal microbial community, or gut microbiota. Utilizing the zebrafish as a model host organism, we documented the development of the gut microbiota through time, and found a significant shift in the composition of the gut microbiota after the onset of adaptive immunity.

This led us to hypothesize that adaptive immunity is an important determinant of gut microbiota composition. We tested this hypothesis using wild type and *rag1*$^{-/-}$ zebrafish, which lack a functional adaptive immune system. Additionally we tested the robustness of the effects of adaptive immunity to dispersal of microbes between immune-compromised and immune-competent genotypes. We found that adaptive immunity had less of an effect on the composition of the gut microbiota than we expected, although there were intriguing differences in the nature of selection imposed when adaptive immunity was present than when it was absent.

Because "westernization", or market-integration, has been associated with

significant changes in the human microbiota and certain health risks, we used similar analyses to those we applied to the zebrafish system to determine whether market-integration alters the filtering effects of inflammation and intestinal helminth parasites on the intestinal microbial community. We found that market-integration increased inter-subject dissimilarity and reduced inter-subject dispersal. Even small changes in the inflammation marker, CRP, were associated with differences in the gut microbiota, but these effects were reduced in the presence of helminth infection, which has been hypothesized to affect the microbiota by reducing inflammation.

In total, this dissertation provides evidence for the nature and importance of host filters of the gut microbiota across two vertebrate species, as well as providing a framework for future studies of the effects of such filters on the assembly of the gut microbiota.

This dissertation includes previously published, and unpublished, co-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: Keaton Daniel Stagaman

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
University of California, Santa Cruz

DEGREES AWARDED:

Doctor of Philosophy, Biology, 2016, University of Oregon
Bachelor of Science, Ecology and Evolution, 2008, University of California, Santa Cruz

AREAS OF SPECIAL INTEREST:

Ecology and Evolution

PROFESSIONAL EXPERIENCE:

Laboratory Technician, University of Idaho, 2008-2009

Research Assistant, University of California, Santa Cruz, 2007-2008

GRANTS, AWARDS, AND HONORS:

Genetics Training Grant, Dissertation Research, Biology Department, 2010

PUBLICATIONS:

Burns AR, Stephens WZ, Stagaman K, Wong S, Rawls JF, Guillemin K, *et al*. (2016). Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME J* **10**: 655–664.

Stephens WZ, Burns AR, Stagaman K, Wong S, Rawls JF, Guillemin K, *et al*. (2016). The composition of the zebrafish intestinal microbial community varies across development. *ISME J* **10**: 644–654.

Stagaman K, Martinez ES, Guillemin K. (2015). Immigrants in immunology: the benefits of lax borders. *Trends Immunol* **36**: 286–289.

Wong S, Stephens WZ, Burns AR, Stagaman K, David LA, Bohannan BJM, *et al*. (2015). Ontogenetic Differences in Dietary Fat Influence Microbiota Assembly in the Zebrafish Gut. *MBio* 6: e00687–15.

Stagaman K, Guillemin K, Milligan-Myhre K. (2014). Tending a complex microbiota requires major immune complexity. *Mol Ecol* **23**: 4679–4681.

Costello EK, Stagaman K, Dethlefsen L, Bohannan BJM, Relman DA. (2012). The Application of Ecological Theory Toward an Understanding of the Human Microbiome. *Science* **336**: 1255–1262.

ACKNOWLEDGMENTS

I would like to thank Karen Guillemin and Brendan Bohannan for their seven years of excellent advice and invaluable support. Special thanks are also due to Zac Stephens and Kat Milligan-Myhre, who as a senior graduate student and post-doc, respectively, were important mentors in my development as a scientist. I would also like to thank Adam Burns, whose expertise in ecological theory will always surpass my own, for his contribution to the experimental design and analysis of much of the work presented in this dissertation. I would finally like to express my gratitude to Rose Sockol for her help in zebrafish husbandry and genotyping, without whom, chapters III and IV would not have been possible. The research reported in this dissertation was supported by the National Institute of General Medical Sciences of the NIH under award numbers R01GM095385 and P50GM098911. Grant P01HD22486 provided support for the Oregon Zebrafish Facility. The content is solely the responsibility of the author and does not necessarily represent the official views of the NIH. The ACISS computational resources were funded by a Major Research Instrumentation grant (Grant No, OCI-0960354) from the National Science Foundation, Office of Cyber Infrastructure.

Dedicated to my wife, Danielle, and my family, without whom I would not have been able to accomplish this degree.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 1

II. THE APPLICATION OF ECOLOGICAL THEORY TOWARDS AN UNDERSTANDING OF THE HUMAN MICROBIOME 3

Ecological Processes Within Humans 4

Metacommunity Theory and the Human Microbiome 10

Postnatal Acquisition and Development of the Human Microbiome 12

Community Assembly Following Disturbance: Antibiotics as a Paradigm 17

Assembly of the Human Microbiome in the Context of Invaders (Pathogens) 20

Translating Ecological Understanding into Clinical Practice 25

III. THE COMPOSITION OF THE ZEBRAFISH INTESTINAL MICROBIAL COMMUNITY VARIES ACROSS DEVELOPMENT 29

Materials and Methods 31

Results 37

Discussion 48

IV. THE ROLE OF ADAPTIVE IMMUNITY AS AN ECOLOGICAL FILTER ON THE GUT MICROBIOTA IN ZEBRAFISH 50

Materials and Methods 54

Results 58

Discussion 68

V. MARKET INTEGRATION, CRP LEVELS, AND HELMINTH INFECTION INTERACT TO SHAPE THE GUT MICROBIOTA OF AN INDIGENOUS ECUADORIAN POPULATION 73

Materials and Methods 76

Chapter Page

Results 79

Discussion 88

VI. CONCLUSION 94

APPENDIX: PRIMER SEQUENCES 97

REFERENCES CITED 98

LIST OF FIGURES

Figure Page

1. Community ecological processes underpinning the formation of human microbiotas 4

2. Alternative community assembly scenarios 8

3. Experimental design and zebrafish development 33

4. Significant changes in diversity of individual zebrafish intestinal communities throughout development 38

5. Major shifts in bacterial taxa throughout development 39

6. Phylogenetic dissimilarity of microbiota from fish and environmental samples 44

7. Experimental design for comparison of wild type and *rag1*$^-$ gut microbiota 54

8. Dissimilarity between gut communities and their tank water 59

9. Phylogenetic alpha-diversity of gut communities 60

10. R^2 values for fit of neutral model to each tank 62

11. Pairwise Sørensen distances between gut communities 63

12. Significant indicator taxa as determined by *LefSe* 65

13. NMDS ordinations of Sørensen distances between gut communities within each neutral model partition 67

14. Style of life metrics by region 76

15. Phylogenetic diversity (PD) by Region, log(CRP+1), and Helminth infection 79

16. Weighted Unifrac ordinations with significant factors and vectors 82

17. Indicator taxa as determined by *LefSe* analysis 83

18. Inter-subject dissimilarity for groups of interest 85

19. Estimated migration rates and neutral model fit for each region 87

LIST OF TABLES

Table Page

1. Genera found in greater than 90% of larval, juvenile or adult intestines 41

2. Results of multiple regressions comparing community dissimilarity with differences in host age, standard length, and sIgM concentrations 46

3. PERMANOVA results of Sørensen-based pair-wise beta-diversity measures...... 64

4. PERMANOVA results of ordination factors by Sørensen for all neutral model partitions 68

5. ANOVA results for PD for region, helminth infection, and CRP levels............... 80

6. PERMANOVA results on Weighted Unifrac distances for factor interactions..... 81

CHAPTER I

INTRODUCTION

Life is dominated by microbes, in terms of both numbers and mass. As such, all eukaryotic life, as far as we know, lives in intimate association with various microbial species. In vertebrate animals, microbes inhabit nearly every surface of the body that makes contact with the outside world: the skin, genitals, mouth, and gastro-intestinal tract. Far from living as quiet neighbors, animals and their resident microbes interact constantly, creating what will from now on be referred to as the host-microbe system. Host-microbe systems are compelling topics of study for a wide array of disciplines. From the perspective of the host, we have learned that the host-associated microbial community, or microbiota, plays a surprisingly integral role in the proper development of tissues such as as the intestinal epithelium and both the adaptive and innate branches of the immune system (Bry *et al.*, 1996; Rawls *et al.*, 2004; Bates *et al.*, 2006; Hooper *et al.*, 2012), demonstrating that host development and microbial interaction are inextricably linked. From the perspective of the microbial community, the host provides an environment unlike any other: one that can respond directly to the behavior of the microbes on both ecological and evolutionary time scales. Not only is there a much stronger interaction between a host and its associated microbial community than say, between soil and its associated microbial community, but because of the aforementioned contributions by the microbes to host development and physiology, the host has a vested interest in shaping the community of microbes that it harbors. As such, hosts have evolved mechanisms for filtering their microbiota to select for species, functions, or behaviors that provide benefits, or prevent harm, to the host.

This dissertation focuses on a particular set of these host filters and how they interact to shape the host-associated intestinal microbial community. The next chapter, published originally in *Science*, provides a more in depth background on the importance of using well-established ecological theory to understand the assembly of host-associated microbial community rather than simply a host-pathogen paradigm. It is this framework, the ecology of host-associated microbial communities, that I apply to the experimental design and analysis in the subsequent chapters. Chapter III, originally published in *ISME J*, reports foundational studies that establish the zebrafish as a model for studying host-microbe system development and provides evidence that the zebrafish host does play a role in filtering its microbiota. Because there are almost innumerable ways in which hosts can filter their gut microbiota, by necessity this dissertation focuses on just a few host factors that were predicted to have strong, relevant effects. Chapters IV and V, both unpublished and coauthored, provide evidence for how these specific host factors—adaptive immunity, inflammation, helminth infection, and market-integration—contribute to host filtering of the gut microbiota in both zebrafish and human subjects. The results from this body of work will contribute to a growing literature that seeks to provide a comprehensive understanding of the assembly and maintenance of host-associated microbial communities.

CHAPTER II

THE APPLICATION OF ECOLOGICAL THEORY TOWARDS AN UNDERSTANDING OF THE HUMAN MICROBIOME

From Costello EK, Stagaman K, Dethlefsen L, Bohannan BJM, Relman DA. (2012). The Application of Ecological Theory Toward an Understanding of the Human Microbiome. *Science* **336**: 1255–1262.

Each human is an assemblage composed not only of somatic cells but also of many symbiotic species. The abundant and diverse microbial members of the assemblage play critical roles in the maintenance of human health by liberating nutrients and/or energy from otherwise inaccessible dietary substrates, promoting differentiation of host tissues, stimulating the immune system, and protecting the host from invasion by pathogens. A number of clinical disorders are associated with alterations in host-associated microbial communities (the ‘microbiota’), including obesity, malnutrition, and a variety of inflammatory diseases of the skin, mouth, and intestinal tract. Thus, the human body can be viewed as an ecosystem, and human health as a product of ecosystem services delivered in part, by the microbiota.

There is growing interest in the use of theoretical methods to study microbial community ecology, and in particular, host-associated microbiota (Mihaljevic, 2012; Prosser *et al.*, 2007). Recent discoveries of unexpected variation in the composition of the microbiome of healthy individuals (Palmer *et al.*, 2007; Ravel *et al.*, 2011; Wu *et al.*, 2011) highlight the importance of identifying the processes that could possibly give rise to such variation. Ecological theory seeks to explain and predict observable phenomena, such as temporal and spatial patterns of diversity.

Here, we explore how community assembly theory could be used to understand the human-associated microbiota and its role in health and disease. We focus on three scenarios relevant to the assembly of the human microbiome: assembly in previously unoccupied habitats (e.g., postnatal development), reassembly following disturbance (e.g., following antibiotic treatment), and assembly in the context of invasion (e.g., by a pathogen).

Ecological processes within humans

The essential building blocks of community assembly theory encompass the processes that create and shape diversity in local assemblages: dispersal, *in situ* diversification, environmental selection, and ecological drift (Vellend, 2010) (**Figure 1**). In addition, coevolution provides another lens through which to view the human-microbial ecosystem (Dethlefsen *et al.*, 2007), although in this review we focus on shorter-term dynamics at the level of individual hosts.

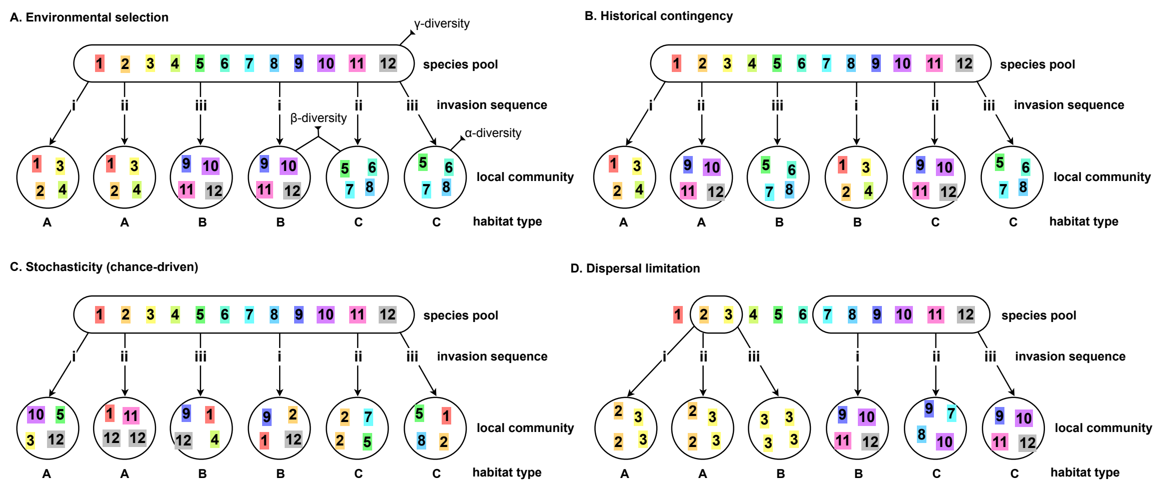


Figure 1. Community ecological processes underpinning the formation of human microbiotas. "Species" (strains, genotypes, or other focal entities) are added via dispersal

and local diversification, and relative abundances are shaped over time by ecological selection, drift, and ongoing dispersal. The figure illustrates a single individual at three separate time points. Circles represent "local" communities, and process examples are shown as follows: Species 2 disperses to the skin surface from an external source (e.g., another individual) and then outcompetes species 3, which goes locally extinct. Species 7 disperses to the urogenital tract from an internal source (the gut) and then recombines (exchanges genes) with species 12 via HGT giving rise to a new genotype (127). In the oral cavity, species 5 increases in abundance but is partially removed by oral hygiene, which returns the community to a more even composition. Species 10 drifts to local extinction in the gut, but is rescued by dispersal from an external source. Species 13, 14, and 15 arrive but are excluded by host or community-level filters.

Dispersal, or the movement of organisms across space, is a fundamental process by which diversity accumulates in local microbial communities. Dispersal tends to emphasize a view of the human body as an "island", a patch of habitat that is continually sampling the pool of available colonists. The list of available colonists may be influenced by microbial traits—those affecting dispersal efficiency, transmission routes, and "ex-host" survivability—and by patterns of host contact and carriage, among other factors. The concept put forth in the late nineteenth and early twentieth centuries that 'everything is everywhere, but the environment selects' had a powerful impact on thinking about community assembly (O'Malley, 2007), but a more recent appreciation of other ecological processes (such as local microbial species diversification) suggests that this conceptualization was overly simplistic. Controlling infectious disease transmission depends on accurate models of host-to-host microbial dispersal (Koopman, 2004), and these could guide investigations into the dissemination of the human microbiome. Selection favors efficient dispersal in pathogens, but perhaps less so among beneficial bacteria, because the host is harmed by the first and not by the second; for beneficial microbes, transmission routes such as direct or close contact may be more important.

Notably, the density and spatial arrangement of host habitat patches has been highly dynamic throughout human history.

A second process operating in microbial communities is local diversification. Unlike in most plant and animal communities, this process can take place over short ecological time-scales for microbes. Large microbial population sizes, high growth rates, and strong selective regimes, all of which can be found in the human body, facilitate rapid microbial adaptation via mutation or recombination. Recombination via horizontal gene transfer may be especially common among members of the human microbiota, especially those sharing the same ecological niche (e.g., body site) (Smillie *et al.*, 2011). Microbial diversification may also be driven by interactions with phage in the human body. Dispersal and diversification may interact (Urban *et al.*, 2008); for example, immigration may suppress adaptive radiations (Fukami *et al.*, 2007).

Relative abundances in local communities are shaped over time by a third process, environmental selection. When considering environmental selection, or niche-based interactions, the human body can be viewed in two ways. First, it can be viewed as a "habitat filter", a collection of resources and conditions allowing the growth of some microbes, but not others, emphasizing the selection of microbial traits that permit survival and growth in the host. In this view, the host shapes the microbiota, but not the other way around. Body temperature is an example of such filtering, since microbes alter body temperature (causing fever) only when they transgress host anatomic boundaries. Second, the human body and its symbionts may be viewed as a community of interacting cells. This view differs from the habitat filter view in that it assumes strong feedbacks between hosts and microbes, and among microbes. This view assumes that the host shapes the

microbiota, and vice versa. Interactions between the host immune system and the microbiota might be best represented by this view (Hooper *et al.*, 2012). Importantly, the overall patterns that arise from dispersal and environmental selection (ecological interactions) can vary as a function of the spatial scale over which these processes occur (Kerr *et al.*, 2002).

In addition to selection-driven changes, species abundances may fluctuate due to the fourth ecological process, known as ecological drift, or demographic stochasticity. As a result of this process, low-abundance species (e.g., recent immigrants, antibiotic sensitive strains, or strains occupying niches with low carrying capacity) are more likely to proceed towards local extinction and become lost from the system, unless they have (or can gain) a competitive advantage, can access a different niche, or become replenished by dispersal from outside the community. Thus, dispersal can effectively "rescue" species from the brink of local extinction, or thereafter.

Finally, the human habitat can be understood as a host-symbiont "holobiont", and as such, an ecological system under selection to minimize conflict between individual members. This view emphasizes the dominant role of co-evolution in the assembly and dynamics of the human ecosystem and reminds us that long- and short-term selective pressures on the human microbiota are not necessarily aligned. Any mutualistic trait that imposes a cost on the microbes that express it – such as producing dedicated molecules to interfere with pathogens or modulate host immune activity – represents a trade-off between the immediate selection against that cost and the long-term selection in favor of mutualism (Dethlefsen *et al.*, 2007).

In summary, different views of humans as microbial habitats make different assumptions about the processes most important to the assembly and dynamics of human microbiome. Community assembly can be conceptualized as being niche-based, dispersal-limited, historically contingent, or random, depending on the relative contributions of habitat conditions, colonist availability, arrival order (and timing), or chance-driven events, respectively, in shaping observed patterns (**Figure 2**). Metacommunity theory integrates the four processes described above and provides a useful framework for considering the community assembly in the human body (Vellend, 2010).

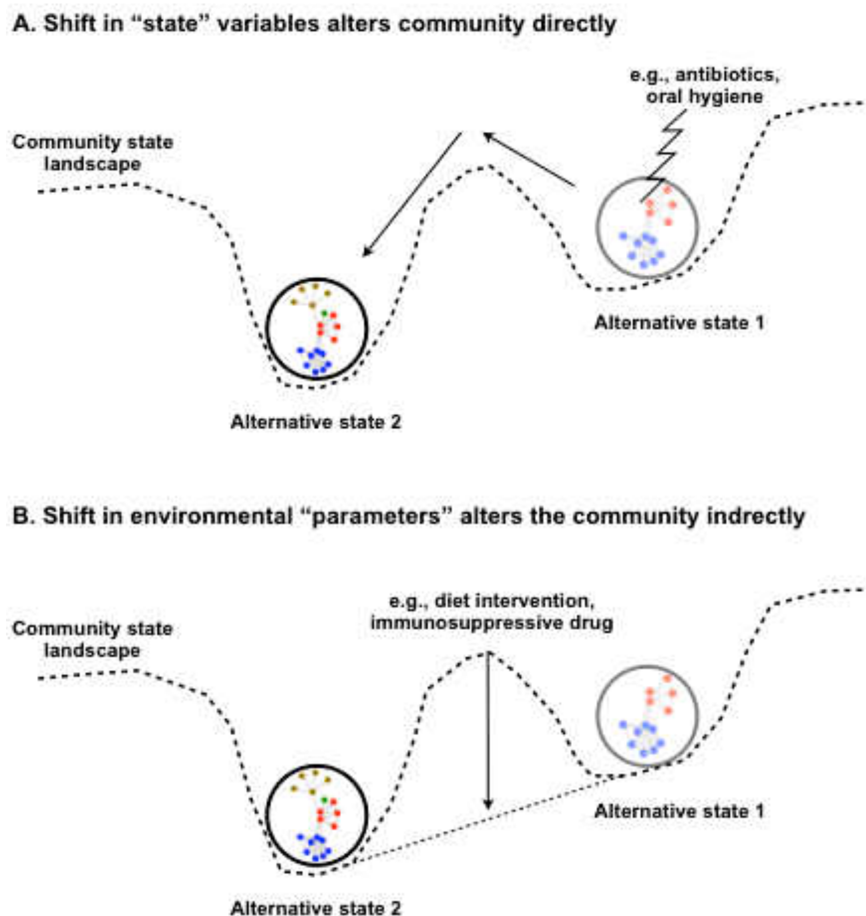


Figure 2. Alternative community assembly scenarios could give rise to the compositional variations observed in the human microbiota. Microbial species are

represented by numbers surrounded by dashed boxes. In A, C, and D, six local communities sample species from a single pool of available colonists. In A, local species composition is driven primarily by environmental selection: regardless of invasion order, habitats with initially similar conditions select for similar assemblages. In C, the opposite is true: regardless of initial habitat conditions, historical contingencies (i.e., differences in the timing and order of species invasions) determine assemblage composition (usually attributed to priority effects). In D, neither habitat nor history matter: local communities assemble via random draws from the species pool (here, like rolling a 12-sided die because all species occur at the same frequency). In B, dispersal barriers cause local communities to assemble from different species pools. For each of the two pools, local communities may assemble as in A, C, or D. Panel A (arbitrarily chosen) illustrates various diversity measures: gamma diversity refers to that of the "regional" species pool, i.e., the total diversity of the local communities potentially connected via dispersal; beta diversity refers to the amount of diversity shared between local communities; and alpha diversity refers to the amount of diversity found within a local community. Panel B demonstrates the relationship between low gamma diversity (the species pool containing only two species, 2 and 3) and low alpha and beta diversity. Although it's likely all four scenarios (A-D) play a role in generating microbiome variation, suggested examples highlight where each might predominate.

Metacommunity theory and the human microbiome

One of the key theoretical frameworks used in studies of community assembly is neutral theory (Hubbell, 2001) in which it is assumed that dispersal, diversification, and ecological drift are purely chance-driven processes. It is a null model because it invokes neither environmental selection nor inherent differences in species' ability to disperse or diversify. Although neutral theory on its own is quite valuable in testing this null hypothesis, an ideal model for the assembly of the human microbiome might accommodate it with alternative theories, combining the strengths of transmission dynamic models (e.g., inclusion of host contact and carriage dynamics) with those from island biogeography and community ecology (e.g., focus on communities, rather than individual pathogens). One such approach is metacommunity theory (Leibold *et al.*, 2004), which could be especially useful for modeling host-associated communities.

Metacommunity theory views the world as a collection of patches, or spatially distinct areas of suitable habitat surrounded by a matrix of unsuitable habitat. These patches each contain a community of organisms, and these spatially distinct communities are connected together to form a metacommunity by the dispersal of organisms from patch to patch. Human populations can be viewed likewise, with host-to-host dispersal linking microbial communities. Metacommunity theory is especially helpful for understanding the relative importance of dispersal and environmental selection in shaping host-associated communities (Mihaljevic, 2012), two issues that have received relatively little attention in studies of the human microbiome.

The predictions of metacommunity theory depend on the frequency and extent of dispersal, differences in the traits of individual organisms, and the degree to which patches vary in their environmental conditions (Leibold *et al.*, 2004; Ellis *et al.*, 2006). Dispersal can be infrequent and localized, or widespread and frequent, as discussed above. In some metacommunity models, patches are assumed to be essentially identical, and that movement among patches determines variation in community membership. Such models might be especially appropriate for populations of closely related hosts. Other models assume that patches vary strongly in their available niches, and that variation in community membership results at least in part from environmental selection (e.g., underpinned by host genetics or diet).

Metacommunity theory enables one to predict the conditions under which community dynamics within a patch are driven by immigration from outside versus local adaptation. In the human microbiome, low dispersal rates favor adaptation within a patch and high dispersal favors immigration. This concept could be useful, for example, in

understanding responses to antibiotic use. If acquisition of antibiotic resistance were primarily a result of immigration, then interventions focused on quarantine and hygiene would be more effective than those focused on altering antibiotic duration or dose (see “An ecological approach to managing invasions” below).

While metacommunity theory has been used to elucidate the drivers of non-host associated microbial community membership and dynamics (Logue *et al.*, 2011; Ofiteru *et al.*, 2010; Van der Gucht *et al.*, 2007), it has rarely been used to study host-associated communities (Hovatter *et al.*, 2011; Sloan *et al.*, 2006), e.g.,, to explore the stringency of host selection and its dependence on the microbial group or the age of the host. Ultimately this information will result in a better understanding of how microbes are “filtered” by the host and, conversely, how microbes evade this filtering. This information is crucial if clinicians are to directly manipulate host-associated communities, through, for example, the design of probiotics capable of evading host filtering and establishing within a host.

The effective application of metacommunity theory (and assembly theory, in general) to the human microbiome requires a preliminary understanding of how the microbiome varies across hosts and over time. In the subsequent discussion, we review our current understanding of this variation, focusing on the dynamics of communities in newly created habitats (e.g., neonatal colonization), the dynamics following disturbance (e.g., after antibiotic treatment) and that following invasion (e.g., by a pathogen). We chose these scenarios because they represent and reveal the fundamental types of assembly relevant to human health.

Postnatal acquisition and development of the human microbiome

Babies are born essentially sterile, and acquire their microbiome from their surroundings. The postnatal assembly of the human microbiota plays an important role in infant health, providing resistance to pathogen invasion, immune stimulation, and other important developmental cues early in life (Mackie *et al.*, 1999b). Acute and chronic disorders, such as necrotizing enterocolitis, antibiotic-associated diarrhea, malnutrition, inflammatory bowel disease, and asthma have been linked to inadequate, inappropriate, or disrupted postnatal microbiome acquisition and development (Murgas Torrazza and Neu, 2011). Mechanisms controlling the appearance of bacteria in healthy infants have been studied for well over a century (Escherich, 1988), and many have likened microbiome development to ecological succession (Mackie *et al.*, 1999b; Savage, 1977; Schaedler, 1965). Succession, as a mode of community assembly, has largely emphasized deterministic processes, being described often as orderly and predictable, but the importance of stochastic and/or historical events has also long been recognized.

In the absence of microbial invasion of the amniotic cavity, which is thought to be a rare, pathologic condition, rupture of membranes signals the moment when microbes, most likely of maternal vaginal origin, first gain access to the neonate. Vaginally delivered infants clearly receive a strong input of vaginal, and possibly other urogenital or fecal microbiota as they pass through and exit the birth canal (Dominguez-Bello *et al.*, 2010; Mändar and Mikelsaar, 1996). Vaginal microbiome composition in non-pregnant, reproductive age women is highly dynamic, and is characterized by at least five compositional classes delineated by different, dominant *Lactobacillus* species, or a lack of *Lactobacillus* dominance. There is frequent class switching over time, including to and

from compositions indicative of bacterial vaginosis, even in the absence of symptoms (Ravel *et al.*, 2011; Brotman *et al.*, 2010; Gajer *et al.*, 2012). Whether these dynamics occur similarly in pregnant and postpartum women has important implications for the initial colonization of vaginally delivered infants; if they do, infant-to-infant variation in the composition of initial colonists may be imposed in some cases by maternal vaginal microbiome class at the time of delivery. Likewise, maternal gut microbiome types (Wu *et al.*, 2011; Arumugam *et al.*, 2011) may also determine the pool of colonists available to vaginally-delivered infants at birth. Thus, variation among neonate microbiomes may reflect variation in maternal microbiomes, but this has not been widely tested for maternal habitats other than the vagina. At the time of delivery, microbiomes do not differ consistently among infant body sites (Dominguez-Bello *et al.*, 2010), implying that sampling is driving initial community assembly, with minimal filtering by the infant host.

Delivery mode also determines microbial exposure at the time of birth. For example, infants delivered by cesarean section do not receive contributions from the vaginal microbiota, and instead, are exposed initially to what appears to be ambient skin microbiota (Dominguez-Bello *et al.*, 2010). Incidental exposures to maternal (or other) gut or vaginal microbiota may occur later in cesarean section infants, at low density or low frequency, and may be inadequate for outcompeting already established strains. For example, cesarean section infants display reduced abundances and/or incidences of colonization by the genera *Bacteroides* and *Bifidobacterium* early in development relative to vaginal deliveries (Bennet and Nord, 1987; Penders *et al.*, 2006). Delivery mode effects can persist for months, and may have consequences for infant health; cesarean section infants have a higher risk for some immune-mediated diseases (Decker *et al.*,

2010; Kuitunen *et al.*, 2009; van Nimwegen *et al.*, 2011). The ambient environment may also play a role in colonization at delivery; infants delivered at home versus the hospital were colonized differently at 1 month of age (Penders *et al.*, 2006). Thus, dispersal limitation imposed by certain medical interventions may contribute to inter-individual variation early in life.

Over the first few months – roughly up until the first solid foods are introduced – a fairly well constrained range of stereotypical bacteria appear in the feces (distal gut), alpha diversity (species richness and evenness within communities) generally increases, and aerobes are thought to be supplanted by facultative and then strict anaerobes (Mackie *et al.*, 1999b). Exclusive breast-feeding has been associated with selection for increased abundance of particular *Bifidobacterium* species whose genome sequences reflect specialized use of human milk oligosaccharides and similar host-derived substrates (Sela *et al.*, 2008), or for other bacteria such as *Bacteroides* that could compete for the same ecological niche (Marcobal *et al.*, 2010). Strikingly, during this early phase, microbiota composition is highly dynamic within and between infants (Palmer *et al.*, 2007; Mackie *et al.*, 1999b; Favier *et al.*, 2002; Koenig *et al.*, 2011; Trosvik *et al.*, 2009), with temporal variation characterized by periods of relative stability (for varying lengths of time) punctuated by abrupt shifts in composition and structure. In some cases, these shifts can be linked with life events that likely impose environmental selection such as fever, formula feeding, or antibiotic therapy (Palmer *et al.*, 2007; Koenig *et al.*, 2011; Savino *et al.*, 2011). Extraordinarily parallel transitions observed in a pair of dizygotic twins suggest that exposures (shared exposures in their case) can also play an important role during this phase, driving within and between infant variation (Palmer *et al.*, 2007). This

finding emphasizes the need to better understand how infants sample their environment over time, e.g., whether outside-of-host environmental reservoirs or direct host exchange paradigms prevail, and with regard to the frequency and extent of dispersal (as discussed above). Abrupt shifts might reflect opportunistic invasions by better-adapted species or subtle filtering by the host. An infant's unique developmental path through this early unstable phase may have longer-term health implications. For example, recent work has shown that colonization during the neonatal period has a particularly important effect on mucosal immune development (Inman *et al.*, 2012; Olszak *et al.*, 2012).

The introduction of solid foods and weaning are associated with the onset of a transition towards an adult-like gut microbiome. Differences due to early exposures such as delivery mode fade as microbiota compositions become more canalized. Life events like illness, diet modification, and antibiotic therapy can still impose disturbances, although specific compositions appear to recover. Taxa characteristic of the adult eventually establish, but the process of microbial community assembly appears to extend past the first year of life and into childhood (Palmer *et al.*, 2007; Koenig *et al.*, 2011). If there is an imprint of microbial flow from parents to children, it is either difficult to detect at early ages and/or else emerges gradually later in life. In one study, fecal patterns of bacterial taxonomic diversity in one year olds were not found to be significantly more similar to those of their parents than to those of unrelated adults (Palmer *et al.*, 2007); but in another study, patterns of microbial diversity in adult twins were slightly more similar to those of their mother (Turnbaugh *et al.*, 2009). These findings suggest that we acquire microbes from competing sources other than, or in addition to, our family members. Further, there may be strong selection for an individualized microbiota. Describing the

adult state as “stable” may not suffice when stability is defined as the permanent coexistence of locally occurring species (Fukami and Nakajima, 2011), because even adult gut composition appears to change slightly over time (Caporaso *et al.*, 2011). In summary, microbiome assembly in newly created habitats likely involves a gradual shift from conditions under the strong influence of dispersal limitation, and stochastic, and/or historical factors, towards conditions increasingly influenced by environmental selection by factors such as diet, with weaning as a strong catalyst, and with development towards adult-like composition continuing into childhood.

Community assembly following disturbance: antibiotics as a paradigm

The assembly of human-associated microbial communities does not, in general, proceed smoothly to a stable climax state which then resists further changes in composition. Disturbances often remove or kill some fraction of the community, providing an opportunity for remaining community members or new colonists to increase in abundance. For example, personal oral hygiene removes bacterial biofilm from teeth, and an antibiotic affects not only the targeted pathogen but also members of the normal microbiota. The former case represents a deliberate attempt to interrupt the development of microbial communities that might be associated with periodontitis, the latter case, an inadvertent consequence of modern medicine. In addition to causing a shift in the community (or state variable), disturbance may also involve a shift in habitat parameters, such as host diet. In many cases, a crucial unknown is resilience – that is, the degree to which the post-disturbance community returns to its former state. While most work on community resilience has considered resilience in terms of community taxonomic

composition, assessment of community function and ecosystem services may be even more important.

The effect of antibiotics on the gut microbiota serves as a paradigm for disturbances in human-associated communities. Antibiotics are now one of the most common and important forms of disturbance of the human microbiota; on any given day, approximately 1-3% of people in the developed world are exposed to pharmacologic doses of antibiotics (Goossens *et al.*, 2005). Over the past several decades there has been increasing concern about the spread of antibiotic resistance among pathogens, as well as growing concern that antibiotic use may disrupt the host-microbe interactions that contribute to human health.

Antibiotic therapy is meant to achieve a sufficient concentration of the drug for a sufficient duration in a particular body site so that the targeted pathogen is eliminated. Even if this aim were always attained, the antibiotic will be found at a range of concentrations at many locations in the body, depending on the mode of administration and its pharmacodynamic properties. Where members of the indigenous microbiota are exposed to antibiotics that affect their growth without killing them, there is selection for resistance. Human gut and oral communities are recognized as reservoirs for the evolution and horizontal transfer of antibiotic resistance determinants, including to pathogens (Smillie *et al.*, 2011; Roberts and Mullany, 2014; Salyers *et al.*, 2004). However, antibiotic resistance among the microbiota is one of several mechanisms that may act to enhance the resilience of the indigenous communities, hence preserving their beneficial ecosystem services. Others may include population-level resistance via stress-

response signaling (Lee *et al.*, 2010), and the existence of dormant persister cells (Vega *et al.*, 2012) or refugium-like locations (e.g., mucus layer).

Only a handful of studies have employed cultivation-independent surveys to examine the consequences of therapeutic doses of antibiotics on the human gut microbiota (Dethlefsen *et al.*, 2008; Dethlefsen and Relman, 2011; Donskey *et al.*, 2003; Jakobsson *et al.*, 2010; Jernberg *et al.*, 2007; Young and Schmidt, 2004). While these studies have examined different antibiotics and employed a range of sampling strategies, durations and analytical approaches, they all have found that antibiotic treatment alters the composition of the gut microbiota, and that the abundance of most taxa begins to return to prior levels within several weeks. However, the studies are also consistent in showing that various taxa recovered to different extents, and that some do not recover over the duration of the study. The antibiotic effect is greater than the routine temporal variability of community composition (Dethlefsen *et al.*, 2008; Dethlefsen and Relman, 2011; Donskey *et al.*, 2003; Jernberg *et al.*, 2007). Some studies have revealed that the composition of strains within a taxon is sometimes altered, even if the overall relative abundance of taxon members returns to pre-antibiotic levels. In both of the studies that involved measurements of the prevalence of antibiotic resistant strains, elevated levels of resistance persisted to the end of the study (Jakobsson *et al.*, 2010; Jernberg *et al.*, 2007).

Overall, research suggests that the human gut microbiota of generally healthy adults is largely, but not entirely, resilient to short courses of antibiotic therapy, while clinical evidence indicates that extended or repeated courses are more likely to result in serious complications such as the invasion and bloom of *Clostridium difficile* (Owens *et al.*, 2008). Perhaps over short courses of antibiotics, a sufficient, although possibly quite

small number of residual cells from most of the large, pre-existing populations survives to recolonize the gut. An increasing number of these residual cells may be lost with longer or repeated courses of antibiotics. Thus, reassembly of the microbial community following extended antibiotic treatment may require colonization from outside the host, a process that would likely be more variable and require a longer period of time than reassembly via the filtering of existing populations in the host. In addition, the microbiome may be highly vulnerable to invasion by, or blooms by pathogens during recovery after disturbance, because resources are in high abundance and resident populations are low. The longer recovery time required following extended antibiotic treatment could lead to a higher probability of invasion by pathogenic strains. One can envision a more enlightened strategy for clinical use of antibiotics that includes pre-treatment estimates of a patient's microbial community resilience, e.g., based on the use of a standardized disturbance and monitoring of key community products, mapping of the ecological adaptive landscape, and assessment of the likelihood for community displacement and adoption of a disadvantageous, altered state. Assessments of elevated risk, or of loss of resilience might then prompt efforts at restoration (Lemon *et al.*, 2012).

Little is known regarding the response of the microbiome to frequent antibiotic use. When disturbances take place of greater magnitude or frequency than that to which a community has had an opportunity to adapt, ecological surprises may occur (Paine *et al.*, 1998). Such frequent disturbances may allow the persistence of microbial taxa that are inferior competitors within a given host, but that are maintained across hosts because they have traits that result in widespread and frequent dispersal, i.e., "fugitive" taxa. Such a

scenario is analogous to the patch dynamics paradigm of metacommunity theory (Leibold *et al.*, 2004).

Assembly of the human microbiome in the context of invaders (pathogens)

It is naïve to consider only the interactions between host and pathogen when predicting the likelihood of microbial disease. The latter, for our purposes, is defined as infection of, and proliferation within a host by a species that, in so doing, elicits some sort of pathology. It may be useful to view the pathogen as an invasive species, and the consequences of invasion as a special case of community assembly. Like invasive species in more traditionally studied settings, whether a species can invade a particular community depends largely on niche opportunities: the filters imposed by the abiotic environment and the resistance of the community to colonization by an exotic species (Shea and Chesson, 2002). A successful invasion involves the dispersal of an invader to a new community, initial colonization, and proliferation, steps influenced by the same processes as community assembly more generally.

The environment created by the host determines the number of potential niche opportunities. The nature of this environment is influenced by a number of conditions, including "abiotic" factors such as oxygen levels, pH, and temperature, as well as the abundance and types of available resources, such as composition of the host's diet (Turnbaugh *et al.*, 2006) and carbon sources provided directly by the host, such as mucosal poly- and oligosaccharides (Sonnenburg *et al.*, 2004). In addition, the host immune system acts as an important environmental filter to limit the spatial extent of the microbiota's available niches. The main functions of the mucosal immune system are to

create an inhospitable buffer zone between the microbiota and the host epithelium, and to minimize the incidence of systemic inflammation that would normally be induced in the face of so many bacterial products (Duerkop *et al.*, 2009; Macpherson *et al.*, 2012; Hooper *et al.*, 2012). The immune system performs these functions through three general mechanisms: i) physical barriers such as the inner mucus layer of the colon and stomach, which is generally impenetrable to bacterial cells (Johansson *et al.*, 2008); ii) antimicrobial peptides and mucosal antibodies in the mucus layer that further hinder bacterial colonization of the epithelium (Duerkop *et al.*, 2009); and iii) innate and adaptive immune responses within the regional lymphatic tissues (Macpherson *et al.*, 2012). Inflammation is an important disturbance that alters the host-associated environment. These three mechanisms, in most healthy hosts, select for bacterial species that do well at or near mucosal surfaces or strong barriers such as the skin. However, host filtering is not the only factor influencing the ability of pathogens to invade the host-microbiota community.

One of the most important roles of the microbiota in mediating host-pathogen interactions is protection of the host from pathogen invasion, or 'colonization resistance' (Macpherson *et al.*, 2012; Sekirov and Finlay, 2009). Protection is achieved through induction of the innate and adaptive branches of the immune system, creating an environment that is unfavorable to pathogens (illustrated by the observation that axenic mice (Macpherson and Harris, 2004) and zebrafish (Kanther and Rawls, 2010) have diminished immune responses and impaired barriers to infection), and through direct competition (or filtering – e.g., by lowering vaginal pH). In this case, pathogens are kept at bay by competition with the microbiota for space and resources. This protective effect

is demonstrated by the increased susceptibility to infection of hosts that have had their microbiota altered by antibiotics, a phenomenon well documented by Miller and Bohnhoff in the early 1960s with *Salmonella* invasion of mice pretreated with antibiotics (Bohnhoff and Miller, 1962). The ability of certain anaerobes to limit the invasion and growth of *Clostridium perfringens* in a diet-dependent manner is an example of competition for resources (Yurdusev *et al.*, 1989). *Bifidobacterium breve* produces an exopolysaccharide (EPS) that protects it from the immune response; this allows it to compete for space and colonize the mouse gut at high loads in both the lumen and at the epithelial surface without inducing inflammation (Fanning *et al.*, 2012). Even if invaders do gain a foothold, the indigenous microbiota can block lethality: in mice, some *B. longum* strains can protect against enterohemorrhagic *E. coli*-mediated death by inhibiting translocation of Shiga toxin from lumen to blood (Fukuda *et al.*, 2011).

By viewing pathogens as invasive species we see that the contexts in which they are able to cause disease are the same as those required for any other species that invades and proliferates in a community. Niche opportunities can result from exploiting novel or overly abundant resources (from the host's food), out-competing a commensal species for the same resource, or perhaps most importantly, exploiting niches left open after a disturbance. The importance of exploiting disturbance is well illustrated by the increasing number of cases of disease caused by *Clostridium difficile* (Kelly and LaMont, 2008). *C. difficile* is a "weedy", both native and exotic species that can rapidly fill niches once they are vacant, but in most cases is eventually removed or kept at low numbers in the absence of a disturbance. *Salmonella enterica* serovar Typhimurium is an example of an exotic invasive that exploits disturbance, but in this case, it also causes the disturbance it

exploits. *S.* Typhimurium expresses many virulence factors that create inflammation in the mammalian intestine. A mutant *S.* Typhimurium strain lacking these virulence factors is unable to invade the gut community and cause disease; however, if inflammation is provided by some other mechanism, otherwise avirulent strains are able to invade the host communities (Stecher *et al.*, 2007). Inflammation likely reduces the abundances of other bacteria that would compete with pro-inflammatory pathogens for space. As one possible mechanism, inflammation causes the intestine to produce tetrathionate which *S.* Typhimurium utilizes as an electron acceptor for respiring ethanolamine, a carbon source that cannot be exploited by other bacteria, thus avoiding competition for nutrients (Thiennimitr *et al.*, 2011). By causing acute inflammation, the pathogen is able to alter the native microbiota and effectively colonize and proliferate.

In contrast to the above examples, a reduction in disturbance frequency can also promote invasion by pathogens, as evident in cases of cystic fibrosis. Patients with cystic fibrosis produce thickened mucus, which inhibits the ability of the cilia to remove foreign material from normally sterile lung airways. This lack of constant removal (i.e., impaired innate immune host filtering mechanism), among other factors, allows for the establishment of bacterial communities that would normally not be able to persist at that site (Klepac-Ceraj *et al.*, 2010).

In summary, predicting the success and outcome of infection by pathogens can be aided by framing the issue as an ecological problem of community assembly. Invasion ecology highlights the importance of niche opportunities as determinants of success of invasion, and the manipulation of which might help in pathogen control and disease prevention. Experimental models using gnotobiotic organisms such as mice and zebrafish

will be helpful in understanding the role of community diversity, as well as the role of particular community members in conferring colonization resistance through indirect inhibition and resource competition. In addition, the frequency and magnitude of disturbance plays a crucial role in facilitating both the colonization by exotic invasives as well as the expansion of native species. Finding ways, through prebiotics, probiotics, or pharmabiotics, to alter pathogen or other bacterial species abundances (or to inhibit their detrimental effects on the host) in a specific manner without causing additional disturbance to the community will be very important for preventing and treating disease caused by invasive species.

Translating ecological understanding into clinical practice

An improved understanding, informed by ecological theory, of how microbiomes assemble could alter clinical practice by changing the perspective clinicians bring to the treatment of infectious disease. The traditional perspective has been to think of the human body as a battleground, on which physicians attack pathogens with increasing force, occasionally having to resort to a scorched earth approach to rid a body of disease. Although this perspective has been very successful for several diseases, it has come at a great cost. Even for those diseases for which it has worked, the collateral damage can be severe. As we have discussed, antibiotics often kill beyond the target organisms (Dethlefsen and Relman, 2011) and can increase the chance of invasion by unwanted organisms (such as *C. difficile* (Kelly and LaMont, 2008)).

The body-as-battleground approach ignores the community context of infectious disease, and does not take into account our increasing knowledge regarding the assembly

of the human microbiome. We suggest that it is time for clinicians to abandon the war metaphor (Lederberg, 2000). Given the ecological parallels between assembly of the human microbiome and assembly of other ecological communities, we suggest that human medicine has more in common with park management, than it does with battlefield strategy. To effectively manage a plant or animal community requires a multipronged approach of habitat restoration, promotion of native species, and targeted removal of invasives. We describe below some examples of how such a human-as-habitat approach might alter clinical practice.

An ecological approach to managing invasions. An understanding of the dominant mechanisms of community assembly could directly alter how clinicians treat infectious disease. Consider, for example, the rise of drug-resistant pathogens during the course of drug treatment. We can consider this a "special case" of community assembly, much as we did invasion by a pathogen (see above). We can ask: what is the relative importance of dispersal, diversification, environmental selection and ecological drift in the successful invasion by this drug-resistant strain? If the source of these strains is primarily through random sampling of the external environment, then the most effective preventative strategy may be quarantine and enhanced hygiene. In contrast, if such strains arise primarily through diversification of resident pathogens, then multi-drug treatment may be more effective (to make successful evolution more difficult). If the drug-resistant strains are already present at the outset of treatment and increase in abundance via environmental selection, then drug cycling may be the most effective treatment to reduce the overall competitive advantage of the resistant strains. If drug-resistant strains establish primarily through ecological drift, then disturbance may be crucial to their

establishment (to free up ecological "space" for their invasion). In this case, reducing disturbance of the resident microbiota may be most effective. In this way, a detailed understanding of the relative importance of different community assembly processes can be used to tailor the treatment of disease.

Health as a product of ecosystem services. Humans benefit from a variety of processes supplied by natural ecosystems. Collectively these benefits are known as ecosystem services (Daily *et al.*, 1997). There is growing evidence that human health is a collective property of the human body and its associated microbiome, and thus could be considered a net effect of ecosystem services. We envision clinical medicine focused on managing the human body and its associated microbiome to preserve these ecosystem services. How might this be accomplished? In general ecology, the management of an ecosystem service requires four basic steps (Allan and Stankey, 2009): i) identification of ecosystem service providers (ESPs; taxa that provide specific ecosystem services) and characterization of their functional roles; ii) determination of how community context influences the function of these providers; iii) assessment of key environmental factors influencing the provision of services; and iv) measurement of the spatial and temporal scales at which these providers and their functions operate. This general framework would work equally well for human health-associated ecosystem services. If studies of the human microbiome were structured around these four priorities, the development of an ecological approach to medicine could be accelerated. Progress has been made in identifying ESPs ("biomarkers"; Lemon *et al.*, 2012); for example, declines in *Faecalibacterium prausnitzii* are associated with inflammatory bowel disease, and this organism may be an ESP for health in the human gut (Sokol *et al.*, 2008).

Adaptive management of the human body. Transitioning clinical practice from the body-as-battleground to the human-as-habitat perspective will require rethinking how one manages the human body. In the management of plant and animal communities, a system-level approach known as "adaptive management" has become popular. This approach is a structured, iterative process of decision-making, one that utilizes system monitoring to continually update management decisions (Allan and Stankey, 2009). It has been successfully used to manage biodiversity in a variety of habitats, including communities in highly disturbed environments impacted by overfishing and by climate change (Allan and Stankey, 2009). For the human body, we envision that this approach would involve monitoring of the microbiome during health, to establish a healthy baseline, with more intensive monitoring during disease and treatment. This will require the development of new diagnostic tools that are both accurate and sufficiently rapid to inform decisions regarding therapeutics (Lemon *et al.*, 2012). Such diagnostics are not yet feasible, but given recent advances in our ability to survey the human microbiome, this possibility is not far in the future, especially if we are able to identify particular components of the human microbiome that contribute disproportionately to the maintenance of human health. An adaptive management approach to clinical medicine is the ultimate in personalized medicine, with treatments tailored to individuals based on diagnostic changes in an individual's microbiome, and continually adjusted through regular monitoring. Such an information-intensive approach, guided by ecological theory, has the potential to revolutionize the treatment of disease.

To achieve the goal of treating intestinal diseases by reliably and predictably altering the human gut microbiota, we must first understand how the gut microbiota

assembles. Model organisms provide experimental systems from which we can sample a large number of hosts whose conditions can be highly controlled. The next chapter focuses on the typical assembly of the intestinal microbiota of the emerging host-microbe model organism, the zebrafish.

CHAPTER III

THE COMPOSITION OF THE ZEBRAFISH INTESTINAL MICROBIAL COMMUNITY VARIES ACROSS DEVELOPMENT

From Stephens WZ, Burns AR, Stagaman K, Wong S, Rawls JF, Guillemin K, et al. (2016). The composition of the zebrafish intestinal microbial community varies across development. *ISME J* **10**: 644–654.

Animal development occurs in a dynamic microbial world. The resulting associations between animals and microbes profoundly influence the maturation of their tissues and the function of adult organs. In particular, the development of the vertebrate digestive tract, which harbors the vast majority of microbial cells in the body, is strongly influenced by the presence and composition of the gut microbiota (Bates *et al.*, 2006; Olszak *et al.*, 2012; Semova *et al.*, 2012; Sommer and Bäckhed, 2013). A comprehensive description of animal development must therefore include not only a catalogue of the birth, specification, and differentiation of the animal cells that comprise the body but also the associated microbial cells (McFall-Ngai *et al.*, 2013). Here we present a comprehensive survey of the intestinal microbiota of a single large sibling group (sibship) of the model vertebrate zebrafish (*Danio rerio*) throughout development.

Our study spanned major milestones in zebrafish development under common laboratory rearing conditions. Zebrafish are fertilized externally, therefore post-embryonic developmental stages are referenced as days post-fertilization (dpf). Zebrafish embryos initially develop in essentially sterile chorions and the larval stage begins when the organism hatches from its chorion and first encounters microbes in its external environment (between 2 and 3 dpf). By the time of hatching, most of the larva's organs have been specified but will continue to grow and mature into the adult structures in

interaction with associated microbes. This includes the maturation of the intestine, which is open to the surrounding environment between 3 and 4 dpf, allowing exposure to microbial colonists (Bates *et al.*, 2006). At approximately 5 dpf the yolk becomes depleted and larval zebrafish begin ingesting food. The development and differentiation of zebrafish continues into adulthood. While juveniles of both sexes have ovary-like gonads, they differentiate into sex-specific gonads by approximately 4 weeks post fertilization and continue to develop secondary sex characteristics well into adulthood (approximately 10 – 12 weeks post fertilization depending on rearing conditions; Uchida *et al.*, 2002). Initially, the ability of the host to defend against microorganisms is limited to innate immune activities, with the adaptive immune system reaching functional maturity around 4 weeks post fertilization. Many of the attributes that make zebrafish an excellent model for studying vertebrate development, such as its early optical transparency, small size, high-fecundity, and availability of genetic and genomic resources (Howe *et al.*, 2013; Phillips and Westerfield, 2014), also lend it to studies of vertebrate host-microbiota interactions. Large numbers of zebrafish can be maintained in a shared and easily sampled aquatic environment, allowing a high degree of biological replication along with information from associated environmental microbial communities.

The large degree of biological replication that is possible with zebrafish is an important advantage in understanding the extensive interpersonal variation observed in vertebrate-associated microbiota (Friswell *et al.*, 2010; The Human Microbiome Project Consortium, 2012; Rogers *et al.*, 2014). Interindividual variation in humans is greatest during early stages of infant colonization and decreases with age, while bacterial

diversity within individuals generally increases from initial colonization at birth, stabilizing around 2 - 3 years of age (Palmer *et al.*, 2007; Yatsunenko *et al.*, 2012; Avershina *et al.*, 2014). During this period, weaning marks a dramatic transition for the developing infant microbiota as dietary change, the removal of maternally provided immunologic factors and loss of breast-feeding derived microbes begins to shift the intestinal microbiota towards an adult-like composition (Bergström *et al.*, 2014). Thus, changes in diet and physiology over animal development are closely intertwined and likely interact to shape developmental changes in the associated microbiota.

In the present study, we exploit the advantages of the zebrafish model system to determine how associated microbial communities change along with key developmental, environmental and dietary transitions of the host. Ours is the largest study to date of vertebrate intestinal microbiota from a single sibship of animals throughout development. We observed stage specific changes in microbiota composition over development. Within each developmental stage there remained extensive inter-individual variation, despite the fact that the hosts belonged to a single sibship and shared the same rearing conditions and environments. Across development, we observed that the intestinal bacterial communities became increasingly different among individual hosts and distinct from the surrounding environment.

Materials and Methods

Experimental design and sample collection

We surveyed the gut microbiota of a pair of adult zebrafish parents and 135 of their offspring reared concurrently under identical environmental conditions at multiple

stages in their development, using high throughput sequencing of the 16S rRNA gene. To reduce potential effects of host genotypic variation, this population consisted solely of offspring from a single mating pair. These siblings were split evenly among four replicate tanks, resulting in 70 fish per tank, and were raised in a manner intended to generally reflect commonly employed zebrafish husbandry practices, including diet and water type, flow rate and frequency of changes (**Figure 3A**, lower portion). We sampled zebrafish and their surrounding tank environment at multiple time points meant to capture important developmental transitions: when the entire intestinal tract is first open and microbial colonization of the lumen first occurs (4 dpf), once fish must rely on ingesting food for nutrition (10 dpf), the maturation of the adaptive immune system (21, 28, and 35 dpf), sexual maturity and dimorphism (75 dpf), and senescence (380 dpf; **Figure 3A**, upper portion). At each time point we sampled the dissected intestines of multiple fish sampled evenly across each of the four replicate tanks, resulting in 20 fish (5 per tank) per time point for ages 4 through 35 dpf, 24 fish (6 per tank, 3 male and 3 female) at 75 dpf, and 18 fish (6 per each of three replicate tanks) at 380 dpf (**Figure 3A**; some samples were later removed due to poor sequencing depth). We also measured the standard length (SL) of each fish as a metric of zebrafish staging and growth (Parichy *et al.*, 2009). To examine the maturation of the adaptive immune system, we measured transcript levels of secreted immunoglobulin M (sIgM) from the carcasses of the 10, 21, 28, 35, and 75 dpf fish (time points spanning the course of immune maturation). Both SL and sIgM transcript levels increased with development (**Figure 3B**). Within a given age there was much greater variation in sIgM transcript levels than in SL, which showed little variation until the last time point.

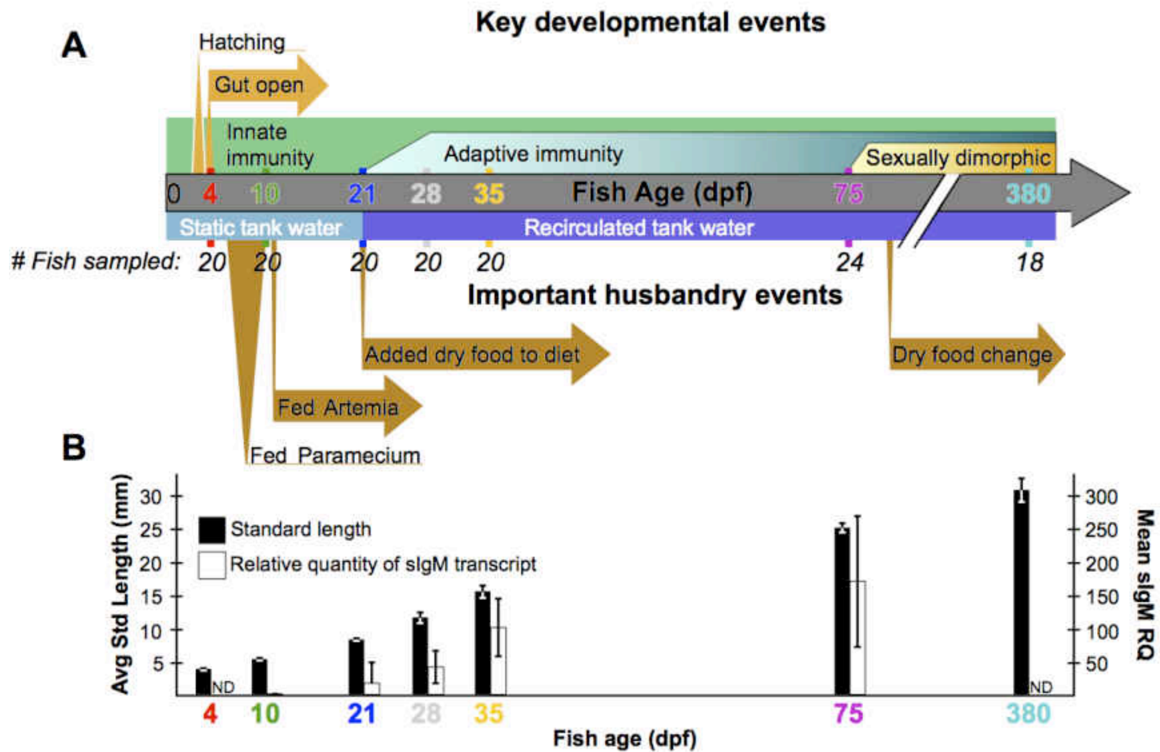


Figure 3. Experimental design and zebrafish development. A) Experimental design showing important developmental events (top) and husbandry events (bottom) during the course of the study. The number of fish initially sampled among 4 tanks at each time point is shown, although post-sequencing rarefaction in some cases reduced this number for analyses (see Materials and Methods). Artemia are commonly called brine shrimp. 4 and 10 dpf fish are considered larvae, 21 – 35 dpf fish are juveniles and 75 and 380 dpf fish are adults. **B)** The mean standard length and secreted IgM (sIgM) transcript levels (a proxy for adaptive immune development) of fish at sampled time points are shown with standard deviations. dpf, days post-fertilization; ND, not determined.

Intestinal and environmental samples were collected and prepared in a manner that minimizes cross-contamination of samples, tanks, and time points. Sampled animals from each time point were collected from the fish facility before they were fed, at approximately the same time of day (between 09:30 and 10:00 AM). Animals were then transported to dissection stations in their own tank water and euthanized by addition of tricaine (2.1 ml of 0.4% tricaine per 50 ml fish water; 0.22 µm filtered) prior to

dissection. Each animal was dissected on a separate, sterile glass slide (larva) or Petri dish cover (juvenile, adults) under a dissecting microscope as previously described (Milligan-Myhre *et al.*, 2011). Larval and juvenile fish dissections were performed using individual-use insect pins, while sterile, individual-use scalpels were used for adult fish. The entire intestine from immediately posterior to the esophagus to the vent was removed intact. The swim bladder and liver were explicitly removed from the intestine, while no effort was made to remove the pancreas (if attached). The intestines were then placed in 2 ml screw cap tubes containing 0.1 mm zirconia-silica beads (Biospec Products, Bartlesville, OK) and 200 (4, 10, 21dpf) or 400 µl (28 dpf and older) of Enzymatic lysis buffer (ELB; Tris-EDTA pH 8.0 with 0.1% v/v Triton X-100; 0.22 µm filter sterilized) prior to freezing in liquid N_2 and subsequent DNA extraction. The remaining carcass (without intestine, swim bladder, liver and likely the pancreas) of each fish was stored in TRIzol (Life Technologies, Carlsbad, CA) at -80ºC for further host RNA extraction and quantification of sIgM transcripts. For each sample age group, DNA extractions were performed on the same day as dissection.

Environmental samples collected at each time point included scrapings from two glass slides each (75mm x 25mm) that were affixed to the bottom (all time points) and sides (28, 35 and 75 dpf time points only) of tanks at the beginning of the study, as well as food samples and 500 ml of water per tank. Water samples were filtered through a 0.2 µm cellulose nitrate filter, which was then exposed to bead beating and DNA extraction from the filter using the same method used for the other samples. Poor DNA extraction efficiency or low number of sequences obtained from several of the environmental samples led to the retention of environmental samples from 4, 10 and 75 dpf time points

only in this study. We also measured multiple environmental parameters from each tank, including temperature, pH, and ammonia, nitrite and nitrate concentrations (data not shown), but these were all found to vary too little over the course of the study to provide explanatory power. Thus, discrete diet changes and a single environmental change from static water in a nursery facility to recirculating water in an adjacent main facility were the only measured environmental variables that appreciably varied between time points.

All zebrafish experiments were conducted in conformity with the Public Health Service Policy on Humane Care and Use of Laboratory Animals using standard protocols approved by the Institutional Animal Care and Use Committees of the University of Oregon and the University of North Carolina at Chapel Hill.

Illumina library preparation and 16S rRNA gene sequence analysis

The microbial communities of individual samples were characterized by Illumina (San Diego, CA) sequencing of 16S rRNA gene amplicons. In order to obtain Illumina compatible amplicons that were amenable to a high degree of multiplexing, we employed a two-step PCR method to add dual indices and Illumina adapter sequences to the V4 region of the bacterial 16S rRNA gene (see Appendix) and obtain paired-end 150 nucleotide reads on the Illumina HiSeq 2000 platform. Illumina sequence reads have been deposited under the NCBI SRA accession number SRP047327.

16S rRNA gene Illumina reads were processed using methods implemented by mothur 1.28.0 (Schloss *et al.*, 2009) and QIIME 1.6.0 (Caporaso *et al.*, 2010). The final OTU table was rarefied to a depth of 4,250 sequences per sample, allowing us to retain all but seven (one each from 21, 28, 35 and 75 dpf groups and three from 380 dpf group

discarded due to low sequence depth) of the originally collected fish intestinal samples in subsequent analyses. Rarefaction curves showed that at this high depth of sampling we were able to sample a large portion of the OTUs (defined using 97% sequence similarity) and diversity present while still retaining a large number of samples within fish of a given age.

Diversity measures and statistical tests

All measures of community diversity and similarity, including OTU richness, phylogenetic diversity, Simpson's index, and unweighted UniFrac distances, were calculated in R (R Core Team, 2015) using the *vegan* (Oksanen *et al.*, 2013), *picante* (Kembel *et al.*, 2010), and *GUniFrac* (Chen *et al.*, 2012) packages. Permutational MANOVA tests were performed using the *adonis* function from the *vegan* package. Phylogenetic diversity was measured as the total shared branch length of OTUs within each community (Faith, 1992). Tests for unimodality were done using Hartigans' dip test for unimodality (Hartigan and Hartigan, 1985). Identification of significant differences in relative abundances in bacterial classes or KEGG functional groups among age groups was accomplished using the Kruskal-Wallis test with the Benjamini-Hochberg FDR correction. Discriminatory analysis of taxonomic groups among zebrafish ages was performed with *LefSe* (Segata *et al.*, 2011).

Results

Zebrafish development is marked by major shifts in the dominant bacterial taxa of the intestinal microbiota

We set out to characterize the zebrafish intestinal microbiota over key developmental time points under standard laboratory rearing conditions, including diet and environment changes during larval and early juvenile stages. Diet and environment were held constant during late juvenile and an early adult stage (75 dpf), while a late adult stage (380 dpf), after a facility diet change, was added to compare adult microbiota of aged fish (**Figure 3**). Over the course of zebrafish development the diversity of observed intestinal microbiota decreased significantly, both in terms of the number of OTUs (**Figure 4A**; $r^2 = 0.19$, $p < 1 \times 10^{-7}$), and phylogenetic diversity (**Figure 4B**; $r^2 = 0.15$, $p < 1 \times 10^{-5}$), with the largest changes occurring between 35 and 75 dpf, during which time diet and environment were held constant but the fish experienced major developmental changes, such as sexual differentiation. The evenness of communities, however, remained relatively constant over host development (**Figure 4C**; $p = 0.288$). We did not observe significant differences in diversity between stages of adult fish. We noticed that the 10 dpf samples appeared to be bimodally distributed with respect to taxa richness, with half of the samples having high richness and half low. We explicitly tested the unimodality of each distribution, and found that only the 10 dpf distribution was significantly non-unimodal (Hartigans' dip statistic D = 0.1336, $p = 0.002$ for 10 dpf samples and $p >> 0.1$ for all other age groups). While there was no significant difference between the means of the standard lengths of the two distributions (two tailed t-test: $p = 0.4159$), we did observe that the community composition of samples belonging to the high richness distribution were significantly more similar to 4 dpf samples than were those belonging to the low richness distribution (measured by comparing pairwise UniFrac distances: $p < 0.001$). This suggests that at approximately 10 dpf the zebrafish

transition from a larval to a juvenile microbiota, and that the fish we sampled were at different stages of this process despite being the same age. One potential explanation for this pattern would be a difference between fish that had begun consuming exogenous food earlier or later.

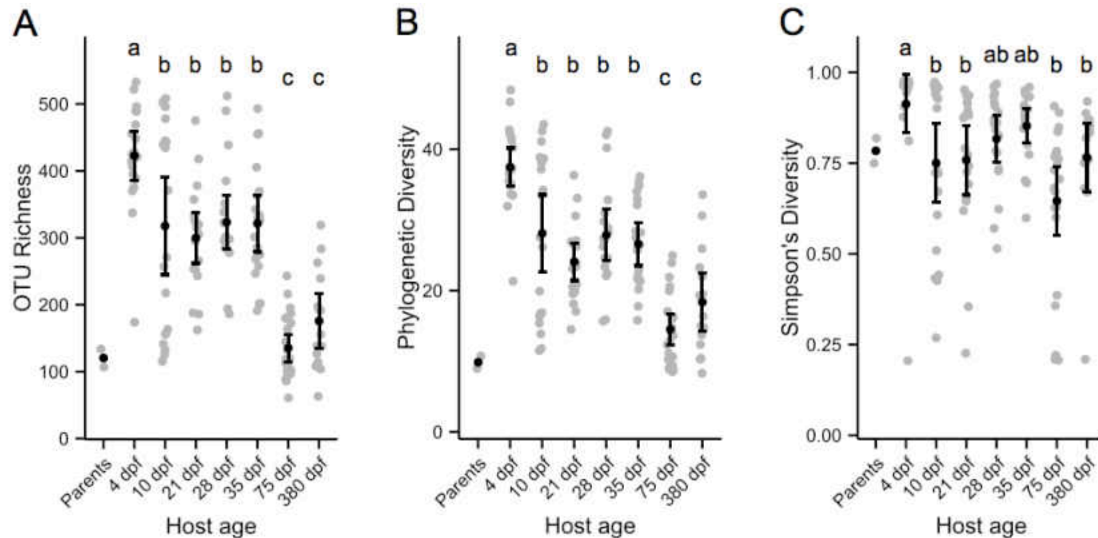


Figure 4. Significant changes in diversity of individual zebrafish intestinal communities throughout development. A) Number of observed taxa. **B**) Faith's phylogenetic diversity. C) Simpson's diversity index. Black circles and error bars represent the means and 95% confidence intervals respectively. Letters above age groups indicate significant differences in the means.

These changes in community diversity were accompanied by significant changes in the phylum-level composition of larval (4 and 10 dpf), juvenile (21, 28 and 35 dpf) and adult (75 and 380 dpf) fish, with particularly large differences in the taxonomic class composition of the Proteobacteria (**Figure 5**). The γ-proteobacteria were the most abundant class of bacteria in the study, and were especially abundant in larval intestines as well as environmental samples. Concurrent with the change in food and environment at 21 dpf, a marked increase in the abundance of α-proteobacteria was observed in the

intestines and was followed by a decrease in abundance during the 28, 35 and 75 dpf age classes, during which time diet and environment were held constant. There was a decrease in the abundance of β-proteobacteria during these same stages, from a peak at 28 dpf. Interestingly, the β-proteobacteria were particularly abundant in all food and environmental samples collected from 10 dpf fish and before (71% average) yet were not consistently as abundant in intestinal samples until 35 dpf, suggesting a time lag between initial exposure and detection of abundant colonization by this class. Although their relative abundances differed in environmental samples, the most abundant β-proteobacteria OTUs in 35 dpf fish were also detected in environmental samples, with the notable exception of an OTU belonging to the family Neisseriaceae that contributed to an average of 5.4% of the reads from 35 dpf fish. The proportion of these Proteobacteria classes decreased strongly in 75 dpf fish, despite being fed the same diet as the juveniles.

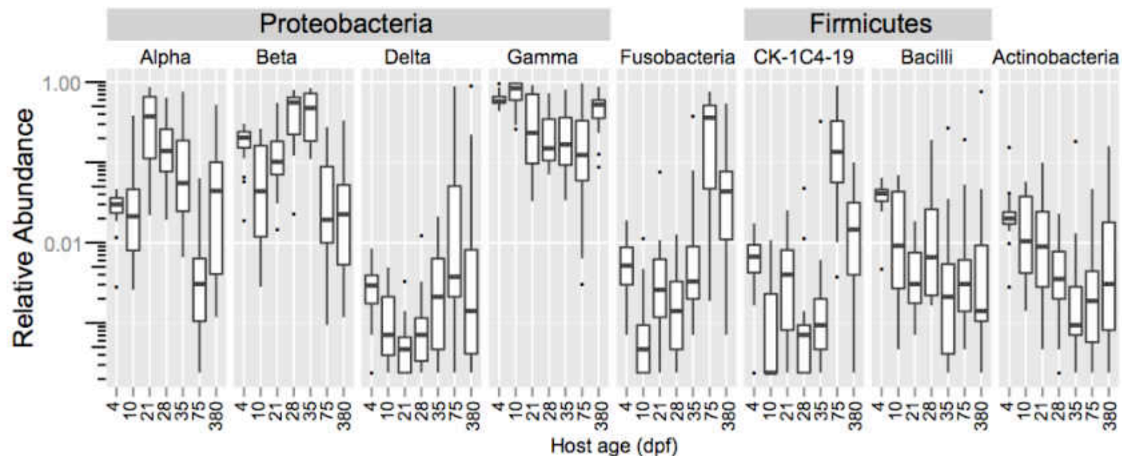


Figure 5. Major shifts in bacterial taxa throughout development. Bacterial classes with >1% average relative abundance across all ages. (All taxonomic classes $p < 0.0001$, Kruskal-Wallis.

Consistent with previous studies of the adult zebrafish intestine (Rawls *et al.*, 2004; Roeselers *et al.*, 2011; Rawls *et al.*, 2006) we found Fusobacteria to be abundant in

the adult stage (75 and 380 dpf) intestinal samples (30% and 12%, respectively), although they accounted for less than 1% of the total community in the 21 - 35 dpf fish. Also in agreement with our previous study (Roeselers *et al.*, 2011) we found a large diversity of Fusobacteria OTUs within intestines (168), with the majority of these OTUs (90%) being classified in the genus *Cetobacterium*, which was found in all 38 adult intestines. OTUs belonging to the Aeromonadaceae family (γ-proteobacteria class) that could not be further classified to genus were the only grouping found in all of the 137 intestinal samples analyzed (**Table 1**). When broken down by developmental group, ten more genera were found in all larval intestines (4 and 10 dpf) and eleven more in all juvenile (21, 28 and 35 dpf) intestines, while only the *Plesiomonas* and *Cetobacterium* genera were also found in all adult intestines. Many of these genera were previously identified as part of a core microbiota of the adult zebrafish intestine (Roeselers *et al.*, 2011). Additional core genera found in greater than 90% of intestines from a given developmental stage group included *Shewanella, Vibrio, Pseudomonas* and *Streptococcus*. The overall abundance of these core taxa varied from nearly 14% for the OTUs within the Aeromonadaceae to less than 1% for the *Streptococcus,* with considerable variation in abundance observed among age groups. An abundant class of uncultured Firmicutes (placed as a separate phylum in some reference taxonomies) referred to as “CK-1C4-19” was found in 89% of all intestinal samples and represented nearly 14% of the total reads in adult intestines. We detected 87 different OTUs from this uncultured class, which represented 3.8% of the total OTUs in the adult zebrafish intestines.

Genus or Best (Taxa Level) Classification	Phylum	All Intestinal Samples		Larval		Juvenile		Adult		Relative Core Genera From Roeselers, et al. 2011 (3)	Zebrafish Isolated Strain Representatives with Genome Sequences (4) (5)
		% Presence in all Intestines (n=137) (1)	% Total Intestinal Reads (2)	% Presence in Larval Intestines (n=40)	% Total Larval Reads (2)	% Presence in Juvenile Intestines (n=57)	% Total Juvenile Reads (2)	% Presence in Adult Intestines (n=38)	% Total Adult Reads (2)		
unclassified Aeromonadaceae (family)	Proteobacteria	100.00	13.92	100.00	14.22	100.00	16.92	100.00	9.84	*Aeromonas*	ZOR0001, ZOR0002
Shewanella	Proteobacteria	97.08	2.65	100.00	6.77	100.00	1.07	89.47	0.76	*Shewanella*	**ZOR0012**
unclassified Enterobacteriaceae (family)	Proteobacteria	95.62	2.82	100.00	7.32	100.00	0.77	84.21	1.32	unclass. Enterobacteriaceae	ZOR0011, ZOR0014
Other Enterobacteriaceae (family)	Proteobacteria	95.62	1.35	100.00	3.46	94.74	0.65	92.11	0.24		
unclassified Comamonadaceae (family)	Proteobacteria	95.62	11.31	97.50	6.03	100.00	21.39	89.47	2.35	*Diaphorobacter*	ZNC0006, ZNC0007, ZNC0008
Plesiomonas	Proteobacteria	95.62	6.81	97.50	8.95	91.23	0.96	100.00	12.35		ZOR0011
Other Gammaproteobacteria (class)	Proteobacteria	94.89	0.88	100.00	2.34	96.49	0.35	86.84	0.18		
Cetobacterium	Fusobacteria	94.89	6.92	85.00	0.38	98.25	1.31	100.00	22.39	*Cetobacterium*	ZWU0022, ZOR0034
unclassified Neisseriaceae (family)	Proteobacteria	94.16	3.76	95.00	0.40	100.00	8.01	89.47	1.14		ZOR0017
Pseudomonas	Proteobacteria	90.51	3.49	100.00	4.13	100.00	4.01	71.05	2.22	*Pseudomonas*	*ZWU0006*
Rhodobacter	Proteobacteria	89.78	1.10	87.50	0.20	100.00	1.51	81.58	1.47		
Other Pseudomonadaceae (family)	Proteobacteria	89.05	0.62	97.50	0.80	94.74	0.60	73.68	0.48		
unclassified CK-1C4-19 (class)	Firmicutes	89.05	4.76	80.00	0.46	91.23	0.94	94.74	13.97		ZOR0006
Other Comamonadaceae (family)	Proteobacteria	86.86	1.10	95.00	0.92	98.25	1.95	65.79	0.07		
unclassified Xanthomonadaceae (family)	Proteobacteria	84.67	0.38	100.00	0.79	78.95	0.21	78.95	0.23	*Stenotrophomonas*	
Vibrio	Proteobacteria	83.21	1.83	100.00	5.39	89.47	0.55	57.89	0.10	*Vibrio*	**ZWU0020, ZOR0035**
unclassified Betaproteobacteria (class)	Proteobacteria	81.75	0.22	97.50	0.39	91.23	0.23	55.26	0.04		
unclassified Rhizobiales (order)	Proteobacteria	81.75	1.44	75.00	0.15	100.00	3.21	63.16	0.23		ZNC0028, ZNC0032
Pseudoalteromonas	Proteobacteria	81.02	2.01	100.00	5.48	96.49	0.89	39.47	0.13		
unclassified Legionellales (order)	Proteobacteria	80.29	0.74	87.50	0.87	98.25	0.85	50.00	0.49		
unclassified Aeromonadales (order)	Proteobacteria	78.83	0.18	97.50	0.19	71.93	0.13	73.68	0.23		ZOR0001, ZOR0002
unclassified Hyphomicrobiaceae (family)	Proteobacteria	78.83	0.30	90.00	0.60	87.72	0.23	57.89	0.11		
Other Rhizobiales (order)	Proteobacteria	78.83	3.18	85.00	1.47	100.00	6.56	42.11	0.07		
Other Betaproteobacteria (class)	Proteobacteria	77.37	0.16	85.00	0.15	100.00	0.26	39.47	0.02		
unclassified Rhodospirillaceae (family)	Proteobacteria	75.91	0.22	92.50	0.36	85.96	0.18	47.37	0.12		
Streptococcus	Firmicutes	72.99	0.63	100.00	1.73	77.19	0.26	42.11	0.08	*Streptococcus*	
Delftia	Proteobacteria	72.26	0.30	92.50	0.33	66.67	0.13	63.16	0.54		ZNC0008
unclassified Pseudomonadaceae (family)	Proteobacteria	70.80	0.41	90.00	0.37	87.72	0.69	28.95	0.07		*ZWU0006*
unclassified Rhodobacteraceae (family)	Proteobacteria	70.07	0.89	40.00	0.03	98.25	1.91	63.16	0.29		
Other Rhodobacteraceae (family)	Proteobacteria	68.61	1.80	20.00	0.02	100.00	3.99	73.68	0.48		
unclassified Vibrionaceae (family)	Proteobacteria	67.88	1.34	72.50	0.23	45.61	0.02	94.74	4.54	*Vibrio*	**ZWU0020**, ZOR0018, **ZOR0035**
Pelomonas	Proteobacteria	66.42	0.81	97.50	1.52	68.42	0.84	34.21	0.04		
Marinomonas	Proteobacteria	65.69	0.25	92.50	0.64	82.46	0.14	15.79	0.01		
unclassified Rhizobiaceae (family)	Proteobacteria	64.23	1.83	47.50	0.05	96.49	4.08	34.21	0.43		ZNC0028
Corynebacterium	Actinobacteria	62.77	0.43	95.00	1.24	56.14	0.12	39.47	0.05		
Mycobacterium	Actinobacteria	62.77	0.26	92.50	0.48	56.14	0.07	44.74	0.33		
Zoogloea	Proteobacteria	61.31	0.48	90.00	1.40	54.39	0.13	42.11	0.08		
Halomonas	Proteobacteria	60.58	1.19	95.00	3.54	70.18	0.35	13.16	0.04		
Erwinia	Proteobacteria	58.39	0.14	95.00	0.38	50.88	0.06	34.21	0.02		
Staphylococcus	Firmicutes	56.93	0.41	92.50	0.64	43.86	0.49	39.47	0.07		*ZWU0021*
Candidatus Rhabdochlamydia	Chlamydiae	50.36	0.21	92.50	0.68	42.11	0.02	21.05	0.03		

1. 'All intestines' includes the 2 parent samples as well as the larval, juvenile and adult intestines from their offspring, listed to the right.
2. Percentages are derived from samples rarefied to 4,250 sequences per sample.
3. Genera shared in all adults from Roeselers, et al. are listed with corresponding genera (or family where genus was ambiguous) from this study. No attempt is made to relate the single core OTU from Roeselers, et al. with only a phylum level assignment.
4. Isolated strains with newly sequenced genomes that belong within a listed taxonomic group. Some strains are listed more than once where they may fall within more than one taxonomic grouping that could not be confidently classified to finer levels with short Illumina reads.
5. Isolated strains which were identified as discriminatory for a given age class are highlighted in **bold** (larval), *italics* (juvenile) or underlined (adult).

Table 1. Genera found in greater than 90% of larval, juvenile or adult intestines.

We next asked if any of these taxonomic groups were strongly associated with particular stages of zebrafish development. For this analysis, we combined the 75 and 380 dpf fish intestinal communities together into a single class ("adult"), and analyzed all of the age classes using a non-parametric test of significance and linear discriminant analysis (LDA) with the defaults implemented by *LefSe* (Kruskal-Wallis; $p < 0.05$ and log 10 LDA score > 2.0). These analyses identified 184 discriminatory taxa, of which the majority (95) distinguished the youngest, non-feeding (4 dpf) age class from all others. The discriminating taxa for the 4 dpf fish largely belonged to the Proteobacteria despite this phylum's abundance in the entire dataset. In order to determine the most highly discriminatory bacterial taxa for each age class we implemented stricter cutoffs for *LefSe* ($p < 0.01$, log 10 LDA score > 3.5). The 10 dpf age class, which had begun feeding on *Paramecium*, were distinguished by the consistent presence of the genus *Mycobacterium*. Notably, this genus contains the fish pathogens *M. marinum* and *M. chelonae*, which were known to be present in our facility during the time of the experiment, and which have recently been shown to be efficiently transmitted to zebrafish via ingestion of infected paramecia (Peterson *et al.*, 2013). While our sequences did not allow us to resolve the species-level identification of these *Mycobacterium* OTUs, we detected *Mycobacterium* sequences in every environmental sample associated with *Paramecium* feeding (8 samples, average 4% abundance), including all three replicates of the *Paramecium* food samples, suggesting a possible transmission route. The juvenile age classes were largely discriminated by the presence of β-proteobacteria lineages, while the adult class was distinguished by the low abundance Bacteroidaceae family, the prevalent Fusobacteria (specifically the *Cetobacterium* genus) and by the CK-1C4-19 candidate

class of Firmicutes. Differences in taxonomic composition throughout development were reflected by differences in the predicted functional capacity of these taxa, which included predicted differences in the representation of genes involved in cell motility and carbohydrate metabolism between adult and younger fish.

Variation in microbial community composition changes over zebrafish development

Despite the genetic similarity and shared environment of their hosts, the overall composition of microbial intestinal communities showed a substantial amount of variation among fish, as measured by the UniFrac distance (**Figure 6**). Communities associated with individual fish were more similar to communities associated with fish of the same age than they were to those associated with fish of different ages (**Figure 6A**; permutational MANOVA; $r^2 = 0.18$; $p < 0.001$). Over time variation among hosts significantly increased, but the effect was small and non-monotonic in the juvenile stages (**Figure 6B**; $r^2 = 0.10$; $p < 0.001$). To determine whether there were possible tank effects, we performed a permutational MANOVA with age, tank, and age by tank as factors. Neither tank nor the interaction of age by tank produced significant effects ($p = 0.930$ and $p = 0.363$ respectively), suggesting there was little to no tank effect that would influence the interpretation of our results. Prior to 75 dpf, we were unable to assign sex to each zebrafish using external traits, however we did not see a significant effect of sex on community similarity at 75 or 380 dpf ($p = 0.11$ and $p = 0.12$ respectively). Given these results, we grouped samples from fish raised in different tanks and fish of both sexes together for the remainder of the analyses.

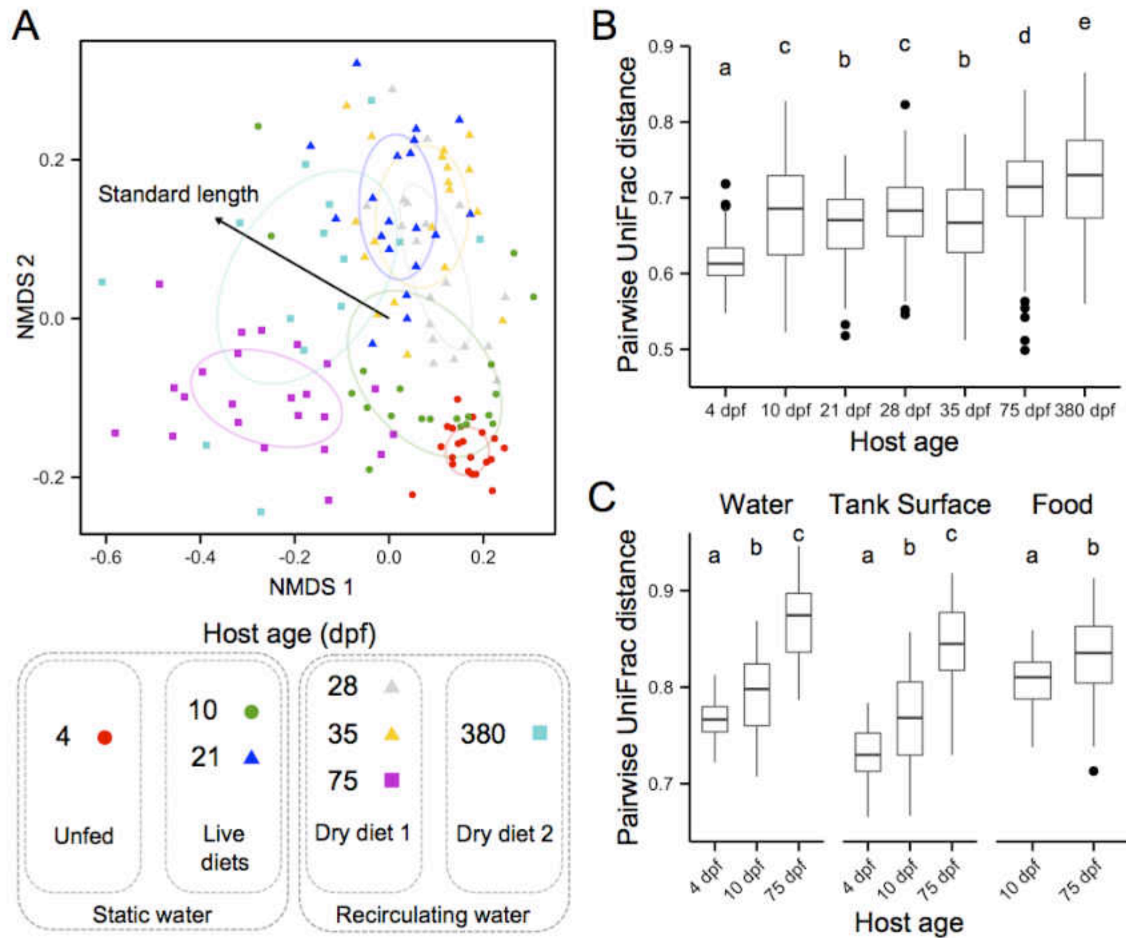


Figure 6. Phylogenetic dissimilarity of microbiota from fish and environmental samples. A) An NMDS ordination of (unweighted) UniFrac distances among zebrafish intestinal communities across development. Host age is differentiated by the color of points, while husbandry conditions (diet and water flow) are described in the legend. The age class of the host is indicated by the shape of points: circles indicate larvae, triangles indicate juveniles, and squares indicate adults. The effect of standard length on the spread of points is shown by a vector. **B)** Pairwise Unifrac distances among fish within each age group. **C)** Pairwise Unifrac distances between fish and environmental communities at each age group. For each boxplot, letters above age groups indicate significant differences in the means.

We next attempted to explain variation in community composition using measures of host age, standard length (SL), and sIgM transcript levels to disentangle the relative influence of time, development, and immune maturation. We used multiple regression analysis (Lichstein, 2007) to partition the variation in pairwise UniFrac distances among

hosts into the total amount of variation explained by the above host variables (i.e. "total"), the variation explained simultaneously by multiple variables (i.e. "shared"), and the variation uniquely explained by each variable independent of the others (i.e. "unique"; **Table 2**). Across the dataset, differences in the standard length of the zebrafish explained more variation in among-host UniFrac distances than did differences in host age, despite host age and standard length themselves being strongly correlated ($r^2 = 0.60$, $p < 0.001$; **Figure 3B**). To determine the potential role of adaptive immune function in structuring communities, as well as isolate the effects of development from changes in diet and housing, we next compared the explanatory power of differences in standard length and differences in sIgM transcript levels (for those samples with measurable sIgM transcription levels and that shared common husbandry conditions; i.e. 28 - 75 dpf samples). We found that standard length was a much stronger predictor than sIgM transcript abundance which explained relatively little variation in UniFrac distances (**Table 2**). It is also worth noting that the explanatory power of standard length was much higher for these age groups, possibly the result of housing conditions being constant for these time points, thus enhancing the relative contribution of host development.

Host Variable	R^2	*P-value**
Across all ages:		
DPF + Standard Length *(total)*	0.2818	0.001
DPF + Standard Length *(shared)*	0.0775	0.001
DPF *(unique)*	0.0057	0.029
Standard Length *(unique)*	0.1986	0.001
For 28, 35, and 75 dpf zebrafish†:		
Standard Length + [sIgM] (total)	0.5084	0.001
Standard Length + [sIgM] (shared)	0.0495	0.001
Standard Length (unique)	0.4427	0.001
[sIgM] (unique)	0.0162	0.037

**P*-values were calculated from a distribution of 1000 random permutations.
†These ages had measurable sIgM transcript levels and shared husbandry conditions

Table 2. Results of multiple regressions comparing community dissimilarity with differences in host age, standard length, and sIgM concentrations.

We computed the pairwise dissimilarity between fish intestinal communities and each of three communities associated with the external environment: the tank water, tank surfaces, and food fed to the fish. Fish intestinal communities were more similar to other fish intestinal communities than they were to any environmental communities (**Figure 6C**; $p < 0.001$ for all comparisons post Bonferroni correction). The dissimilarity between intestinal communities and environmental communities increased over time, ($p < 0.0001$; $r^2 = 0.59$, 0.52, and 0.13 for comparisons to tank water, surfaces, and food samples respectively). As a result, the intestinal communities associated with young 4 and 10 dpf fish were significantly more similar to surrounding environmental communities than were older 75 dpf fish ($p < 0.001$ for tank water, surface, and food environments). This pattern was further manifested by increased differentiation of predicted fish associated metagenomes from predicted environmental metagenomes.

Discussion

The microbial community of the animal gut has been described as an additional host "organ", however its assembly, analogous to the process of organ development, is poorly understood. Here we show that the microbiota of a single sibship of zebrafish exhibits a characteristic developmental trajectory, but the cellular composition is much less stereotyped than developing host tissue. Instead, we observe extensive inter-individual variation in intestinal microbiota composition at each developmental stage, despite our ability to control host genotype and environment, that mirrors the inter-individual variation routinely observed in other vertebrate hosts, including humans (Caporaso *et al.*, 2011; The Human Microbiome Project Consortium, 2012) and mice (Benson *et al.*, 2010; Rogers *et al.*, 2014). We conclude that inter-individual variation in microbiota is a characteristic of vertebrates across development.

This study was designed to understand developmental stage specific intestinal microbiota composition and diversity under standard laboratory rearing conditions of the zebrafish, and provides a reference for future studies investigating the crosstalk between developing zebrafish hosts and their microbiota. We observed major compositional shifts both during periods of development when diet and environment were also changing (i.e. from larval to juvenile stages) as well as when diet and environment remained constant (i.e. from late juvenile to adult stages) suggesting that host physiological development likely has significant effects on the microbiota independent of the other factors. The appearance of an adult-like microbiota in mammals begins shortly after weaning, when the introduction of solid foods and the removal of maternally provided immune factors (in particular breast-milk derived immunoglobulins) impact the composition of the infant

gut microbiota (Rogier *et al.*, 2014; Bergström *et al.*, 2014). In this study we detected bacterial taxa characteristic of adult zebrafish (such as Fusobacteria and the CK-1C4-19 class) early in development but they remained low until the adult stages. While we observed increased sIgM levels during the transition from juvenile to adulthood that could influence these compositional shifts, differences in sIgM levels explained far less variation among microbiota than did standard length. This further supports the notion that morphological changes during development are likely the dominant drivers of changes in the microbiota, at least during periods when diet and environment are constant.

We found that the communities associated with larval fish were more similar to communities associated with the surrounding environment than were adult fish, indicating a greater role of environmental exposure early in development. This is consistent with observed correlations in humans between birth delivery mode and the composition of the intestinal microbiota (Dominguez-Bello *et al.*, 2010), and the relative instability of the intestinal microbiota between early stages of post-natal development in human newborns (Koenig *et al.*, 2011; Mackie *et al.*, 1999a; Bäckhed *et al.*, 2015; Palmer *et al.*, 2007) and in mice (Pantoja-Feliciano *et al.*, 2013). These similarities suggest the intestinal environments of mammals and fish may be similar in some key aspects (e.g. relative changes in oxygen concentration through development), and that environmental exposures including diet can have significant impacts on the observed composition of intestinal microbiota early in development.

The changes we observed in community composition during development highlight the need for careful consideration of developmental context in studies of host-microbe interactions. At the very least, comparisons across studies should strive to use

consistent ages and development stages of the sampled hosts. It is well established that the presence and composition of the microbiota influences a wide array of host developmental and physiologic processes in zebrafish and other animal hosts (Bates *et al.*, 2006; McFall-Ngai *et al.*, 2013; Rawls *et al.*, 2004). Our results highlight observations that some of the observed phenotypic variation in animal studies is due in part to variations in the microbiota. For example, it was recently shown that differences in microbial community composition in wild-type mice alter intestinal IgA levels, thereby differentially influencing susceptibility to a chemically-induced model of colitis (Moon *et al.*, 2015). Likewise, it is possible that our observation of greater variation in sIgM transcript abundance in older fish, with little variation in size, is simply a reflection of increasing microbiota variation between individuals in response to widely varying microbial communities. In the future, it may be helpful to develop and deploy engineered communities of cultured microbes to provide reproducible microbiotas for broad use, or use experimental design strategies that control for microbial variation between individuals, stages, clutches, tanks, pedigrees, and facilities. Our characterization of intestinal microbiota dynamics across zebrafish development and the genomes of representative members of these communities provide a useful resource and framework for such future studies.

Employing this framework as a basis to control for variation between individuals, and tanks, we extended it to include questions regarding the effect of host genotype on the microbiota. Specifically, we test the contribution of adaptive immunity to the filtering of the gut microbiota by the host.

CHAPTER IV

THE ROLE OF ADAPTIVE IMMUNITY AS AN ECOLOGICAL FILTER ON THE GUT MICROBIOTA IN ZEBRAFISH

The chapter is coauthored by Adam Burns, Karen Guillemin, and Brendan Bohannan. I collaborated with all authors in designing the following experiments. Adam Burns contributed equally to sample collection and DNA extraction. I was responsible for sequencing preparation and analysis. Adam Burns contributed invaluable input for the neutral modeling, both in initial interpretation and some of the following text, but all text was primarily written by me. All coauthors edited the text of this chapter.

Adaptive immunity is a fascinating biological phenomenon that only emerged in recent evolutionary history in jawed vertebrates. Its defining characteristics are the ability to modify a receptor repertoire during the course of an individual's life, and to "remember" encounters with previous microbes, making it much easier to prevent or clear subsequent infections by the same infectious agent (Flajnik and Kasahara, 2010). Its role in infectious disease resistance is therefore deeply studied and well characterized. However the role of adaptive immunity in shaping commensal microbial communities is not as well understood. An emerging conceptual model for this role is that of an "ecological filter". The concept of an ecological filter is common to plant ecology and can be defined in that context as "a sieve that filters out individuals from an initial seed population during successive life stages and ultimately determines the success of a seed in producing more seeds" (George and Bazzaz, 1999). In the context of host-microbe associations we can view the host as the environment imposing ecological filters, and the microbial communities as the seeds undergoing this filtering. Given the benefits provided

to the host by certain bacterial taxa and the desirability of selecting against pathogenic microorganisms, it is in the host's interest to filter which microbial taxa comprise its microbiota. Many host factors could act as ecological filters on the microbiota, from literal physical barriers such as intestinal mucus (Johansson *et al.*, 2008), to more abstract filters such as diet (Turnbaugh *et al.*, 2008; Wong *et al.*, 2015). These host factors may be redundant, or may interact in either additive or non-additive ways. Here we explore the specific contribution of the adaptive immune system to filtering the intestinal microbial community in zebrafish.

There is a growing body of evidence that animal immune systems can act as an ecological filter of the microbiota (Hooper *et al.*, 2012), in both humans and in models such as mice. However, such studies, particularly those in mice, often utilize small sample sizes, do not fully quantify differences in communities among genotypes, or involve designs that cannot distinguish genotype-specific effects from those associated with isolation of genotypes in distinct cages (Kawamoto *et al.*, 2014; Shen *et al.*, 2014; Zhang *et al.*, 2015). Only one study (Dimitriu *et al.*, 2013) has looked at the strength of filtering by the immune system relative to transmission of microbes (or possibly other host factors) between cage mates. This study compared the gut microbiota between wild type and immunocompromised hosts that were either separated by genotype or cohoused. They found that when genotypes were cohoused, wild type gut communities became more like immunocompromised gut communities, implying that the filtering effects of adaptive immunity can be overwhelmed by migration from other host communities. However, without replicating treatments, the possibility of differences in gut bacterial communities arising in part from cage effects is still present in such a study. Human

studies that show differences in host-associated communities of the skin (Oh *et al.*, 2013) or expansion of certain potential pathogens in the gut communities of immunocompromised patients (Daniels *et al.*, 2007; Gori *et al.*, 2008), utilize larger sample sizes than many mouse studies, but lack the control over the environmental conditions and genetic backgrounds of subjects provided by model organisms.

The zebrafish is an excellent model for testing the filtering effect of adaptive immunity because many of the limitations of past mouse and human studies can be simultaneously addressed. The zebrafish possesses an adaptive immune system very similar to that of mammals (Rauta *et al.*, 2012), a large number of individuals can be used in a single experiment, and housing conditions can be manipulated easily. Although innate immunity is fully functional at the time of hatching, adaptive immunity in zebrafish does not become fully functional until between 21 and 28 DPF (days post fertilization; Lam *et al.*, 2004; Stephens *et al.*, 2016), allowing experimenters to more easily disentangle the unique contributions of each type of immunity. Recently, we documented how the composition of the gut microbiota changes throughout zebrafish development (Stephens *et al.*, 2016). There were significant changes in microbiota composition after 21 DPF, despite constant housing conditions and diet, which implicates adaptive immunity as a possible ecological filter shaping the gut microbiota. In addition, we inferred the relative roles of neutral and selective processes in shaping the zebrafish gut microbiota over development, by fitting to the data an ecological model that assumes only neutral community assembly processes (such as dispersal and stochastic loss of individuals). We found a trend of decreasing neutral model fit across development, implying that ecological filters become more important in shaping the gut microbiota as

the zebrafish develops (Burns *et al.*, 2016). From these results, we expect that one of the ways that hosts increase selection on their gut microbiota through time is via the maturation of the adaptive immune system.

In order to determine if adaptive immunity is an important ecological filter of the gut microbiota in the adult zebrafish, we compared the gut bacterial communities of wild type (adaptive immunity present) and *rag1*$^{-}$ (adaptive immunity inactive) hosts (Wienholds, 2002). We had three main hypotheses regarding adaptive immunity's role as a filter. (1) Adaptive immunity acts as a filter of microorganisms in the host's environment, resulting in a greater difference between environmental water communities and the intestinal microbiota of hosts with a functional immune system versus those without. (2) The filtering effects of adaptive immunity will lead to distinct differences between the composition of wild type and immune deficient hosts. (3) Adaptive immunity, due to the somatic recombination of B and T cell receptors, has an individualizing effect on the gut microbiota of each host, resulting in greater variation in community composition among wild type hosts than *rag1*$^{-}$ hosts. Because of the potential for transmission of microorganisms and/or host factors among hosts in a shared environment, we suspected that the ability of the adaptive immune system to act as an ecological filter would depend on types and immune state of fish that are exposed to one another. We therefore created a treatment with either a low or high potential for transmission between wild type and *rag1*$^{-}$ hosts by isolating or cohousing genotypes, respectively (**Figure 7**). We hypothesized that increasing the potential for transmission between wild type and *rag1*$^{-}$ hosts by cohousing genotypes would mitigate the effects of adaptive immunity in wild type hosts by increasing the migration of microbial taxa

between hosts and overwhelming the filtering effects of adaptive immunity, making their communities more similar to *rag1*$^{-}$ communities, as demonstrated in mice by Dimitriu *et al.* (2013). In order to measure stochastic tank-specific effects (analogous to cage effects in mouse studies), we maintained three tanks of each housing by genotype treatment. In this way, any significant differences between genotypes would have to be greater than any variance introduced by each tank not related to the presence/absence of adaptive immunity or isolating/cohousing genotypes.

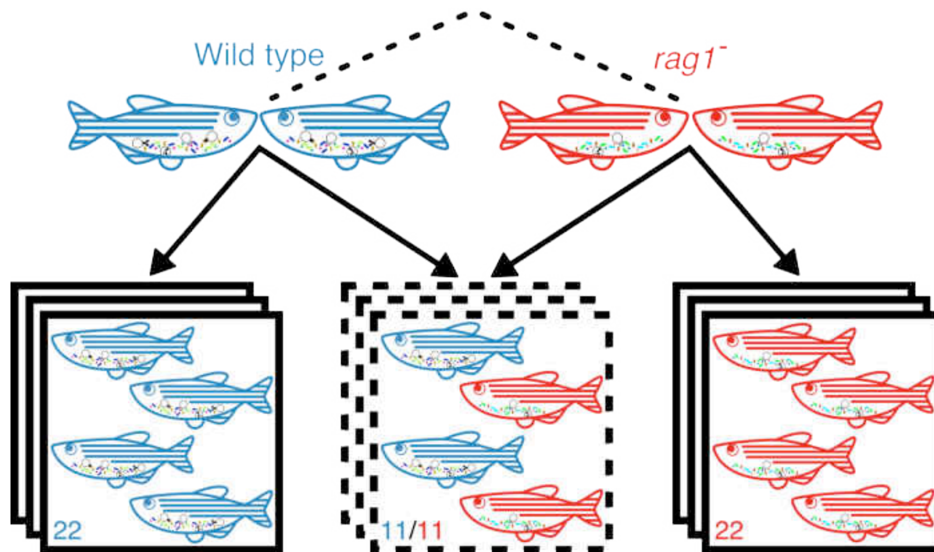


Figure 7. Experimental design for comparison of wild type and *rag1*$^{-}$ gut microbiota. In order to avoid generating 50% heterozygous offspring we cross two set of parents: two *rag1*+/+ parents and two *rag1*-/- parents. These sets of parents were generated from the same line, maintained as heterozygotes. We housed each genotype separately and together, and each housing treatment was replicated three times.

Materials and Methods

Experimental design and sample collection

We crossed two pairs of fish derived from the same heterozygous (*AB/Tübingen* background) line; one pair were both *rag1*$^{+}$ (wild type) and the other pair were both *rag1*$^{-}$. We raised nine tanks of 22 fish each: three tanks of only wild type, three tanks of only

rag1^{-}, and three tanks with equal numbers of both genotypes. At 9 DPF, six fish from each tank were removed for a separate experiment. At 75 DPF, around a month after adaptive immunity should have become fully functional in all wild type fish, all the remaining samples (134) were sacrificed for gut dissection and bacterial DNA extraction. After the fish were removed from the tank, the entire water volume of each tank was run through a filter. Each filter was then subjected to DNA extraction. Water and zebrafish sample handling during collection, dissection, and DNA extraction were all performed as previously described (Stephens *et al.*, 2016).

Carcasses of all fish were also kept after dissection, stored in TRIzol (Life Technologies, Carlsbad, CA, USA) at -80 °C. A subset of the isolated wild type and *rag1*$^{-}$ samples were genotyped, via PCR (Wienholds, 2002), to confirm their presumed identity, and all cohoused samples were genotyped to determine their identity. Five cohoused samples with ambiguous genotyping results (appeared to be heterozygotes) were dropped from analysis after sequencing.

Illumina library preparation and 16S rRNA gene sequence analysis

We characterized the microbial communities of individual samples via Illumina (San Diego, CA, USA) sequencing of 16S rRNA gene amplicons. To prepare amplicons for Illumina sequencing, we used a single-step PCR method to add dual indices and adapter sequences to the V4 region of the bacterial 16S rRNA gene and generate paired-end 250 nucleotide reads on the Illumina HiSeq 2000 platform.

The 16S rRNA gene Illumina reads were processed using methods implemented by FLASH (Magoc and Salzberg, 2011), the FASTX Toolkit (Hannon Lab, 2010), and

the USEARCH pipeline (Edgar, 2010) as detailed in the Supplementary Information. Operational taxonomic units (OTUs) were defined using 97% sequence similarity. Read assembly, quality control, and OTU table building were done on the University of Oregon ACISS cluster, and all subsequent data processing and diversity analysis were done in *R* (R Core Team, 2015).

Diversity measures and statistical tests

Host samples with fewer than 10 000 total reads were removed from analysis, and OTU abundances of the remaining samples were variance-stabilized using *phyloseq* (McMurdie and Holmes, 2013) and *DESeq2* (Love *et al.*, 2014) as recommended by McMurdie and Holmes (2014). Because total reads across water samples were wildly variable, we made community comparisons with the water sample count variance-stabilized, as above, and by rarefying the water samples to the lowest total read count, 29. Despite rarefying to such a lower read count, qualitative differences in community dissimilarity between gut and water communities between genotypes were maintained by both methods. Phylogenetic diversity was measured using Faith's PD (Faith, 1992) as implemented in the *picante* package (Kembel *et al.*, 2010). Expected PD was calculated for each gut community using the *variance.pd* function, also from the *picante* package. The input for the function was the microbial phylogeny for each genotype-by-housing treatment. The function calculated Faith's PD for all subsets of the treatment phylogeny (from full phylogeny to a subset of 1 taxon). The expected PD for each gut community was therefore the *variance.pd* calculated Faith's PD for a subset of the phylogeny equal to the number of taxa in the given gut community. The Z-scores were calculated by

subtracting the observed PD from the expected PD and dividing by the standard deviation of those differences within each treatment using the *scale* function. The z-scores were not re-centered to zero. Sørensen distances (Sørensen, 1948) between communities were calculated and NMDS ordinations performed using the *phyloseq* package. Permutational multivariate analysis of variance (PERMANOVA) tests on Sørensen distances were performed using the *adonis* function from the *vegan* package (Oksanen *et al.*, 2016). For non-distance data, analysis of variance and *post hoc* tests of significance were performed using the *aov* and *TukeyHSD* functions, respectively, from the *R* base packages when the number of observations in each treatment were similar. When sample sizes were not similar, analysis of variance with type III sum of squares was performed using the *Anova* function from the *car* package (Fox and Weisberg, 2011), and *post hoc* tests were performed using the *DTK.test* function from the *DTK* package (Lau, 2013). Diversity data visualization was done with the *ggplot2* package (Wickham, 2009). Analysis of biomarkers (indicator taxa) was done with *LefSe* (Segata *et al.*, 2011) in *Galaxy* (Goecks *et al.*, 2010; Blankenberg *et al.*, 2010; Giardine, 2005).

In order to better understand the ecological processes involved in ecological filtering via the host adaptive immune system, we fit our microbiota data to an ecological model that assumes only neutral community assembly processes (such as dispersal and stochastic loss of individuals). In brief, we applied the Sloan Neutral Community Model for Prokaryotes to the distribution of bacterial taxa in our data and assessed its fit (Sloan *et al.*, 2006). The model predicts that taxa with a high abundance in a source pool of potential bacterial colonists will be found in a greater fraction of those hosts because, being more abundant, they are more likely to disperse by chance, while taxa of low

average abundance are more likely to be lost from individual hosts due to stochastic processes and a lower likelihood of random dispersal. Since we were primarily interested in determining how adaptive immunity mediates the filtering of the microbiota relative to hosts in the population lacking adaptive immunity, we defined the source pool by the average abundance of bacterial taxa across all hosts of both genotypes and housing treatments. We inferred poor model fits to be indicative of non-neutral processes, such as host filtering, being important in distinguishing genotypes and housing treatments from one another. Taxa whose frequency of occurrence across fish were inconsistent with the neutral model, falling outside the 95% confidence interval of the model prediction, were inferred to be subject to host filtering. For example, taxa that are "under represented", that is present in fewer hosts than expected from their mean abundance, could be potential pathogens: successfully prevented from colonizing most hosts but achieving high abundance once established. Taxa that are "over represented", that is found in more hosts than expected from their mean abundance, could be beneficial taxa actively promoted by the host. Using these cutoffs from the model, we partitioned communities into taxa with neutral and non-neutral distributions and determine whether the composition of these partitions differed by host genotype and housing conditions.

Results

Adaptive immunity intensifies host filtering

We hypothesized that adaptive immunity is an important ecological filter on the gut microbiota, with the result that the gut microbiota are more different from the environmental bacterial community in wild type hosts than *rag1*$^{-}$ hosts. To test this

hypothesis, we compared the average beta-diversity, i.e. differences in the abundance and number of taxa between communities, between the gut microbiota of individual hosts and the bacterial communities of their tank water. As shown in **Figure 8**, we found that when adaptive immunity was present (wild type), the gut microbiota was less similar from the environmental community than when it was absent (*rag1*⁻). We also hypothesized that cohousing wild type and *rag1*⁻ zebrafish would allow greater transmission of gut microbes and/or host factors between hosts of differing genotypes, which would either overwhelm the filtering effect of adaptive immunity, or cause the *rag1*⁻ gut communities to appear more filtered. We found that when genotypes are cohoused, wild type gut communities become more similar to the water bacterial communities while *rag1*⁻ gut communities remained unchanged in their similarity to water communities across housing treatments (**Figure 8**).

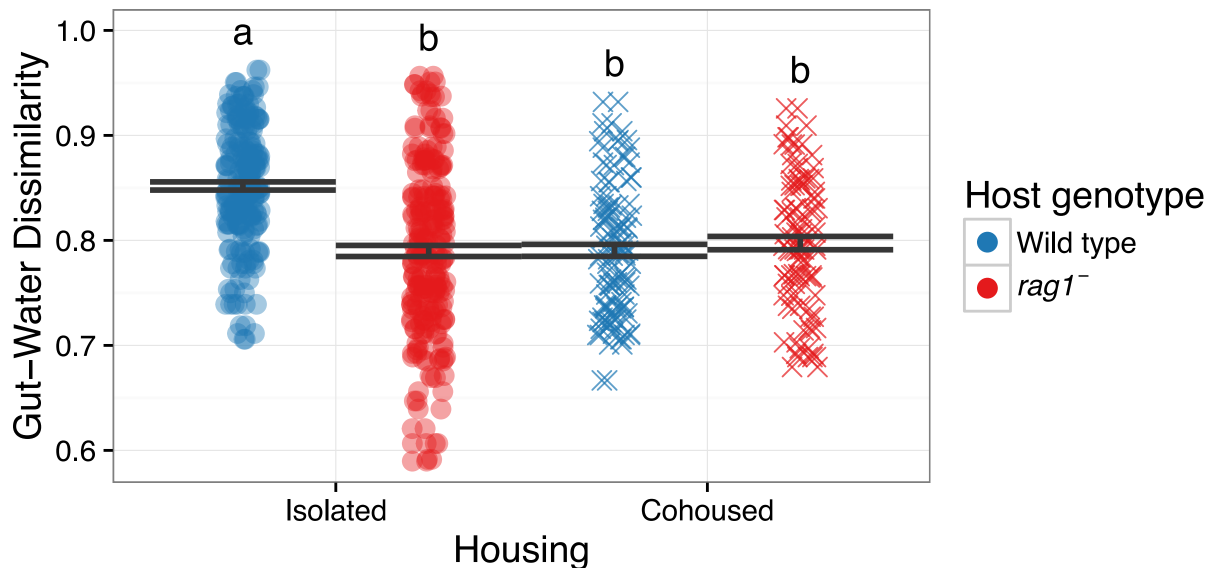


Figure 8. Dissimilarity between gut communities and their tank water. We measured the Sørensen distance between each zebrafish gut community and its surrounding tank water and compared the averages of these distances between each experimental treatment. Shared significance letters indicates no significant difference in means according to the *post hoc* test.

The absence of some or all adaptive immune function has been associated with changes in gut microbiota alpha-diversity, i.e. number and abundance of different taxa within a community, in some mouse studies, though there are conflicting reports as to the direction of this change (Kawamoto *et al.*, 2014; Shen *et al.*, 2014; Zhang *et al.*, 2015), and other studies report no or minimal effects of adaptive immunity on alpha-diversity (Dimitriu *et al.*, 2013; Thoene-Reineke *et al.*, 2014). Without any strong pre-conceived notions, we tested the hypothesized that adaptive immunity in zebrafish alters the alpha-diversity of the gut microbiota. We found no statistically significant differences between genotypes in either housing condition with regard to alpha-diversity (**Figure 9A**). We did find, however, that cohousing *rag1*$^{-}$ with wild type hosts caused a significant decrease observed PD relative to expected PD, i.e. phylogenetic clustering increased significantly for *rag1*$^{-}$ hosts when they were cohoused with wild type hosts (**Figure 9B**).

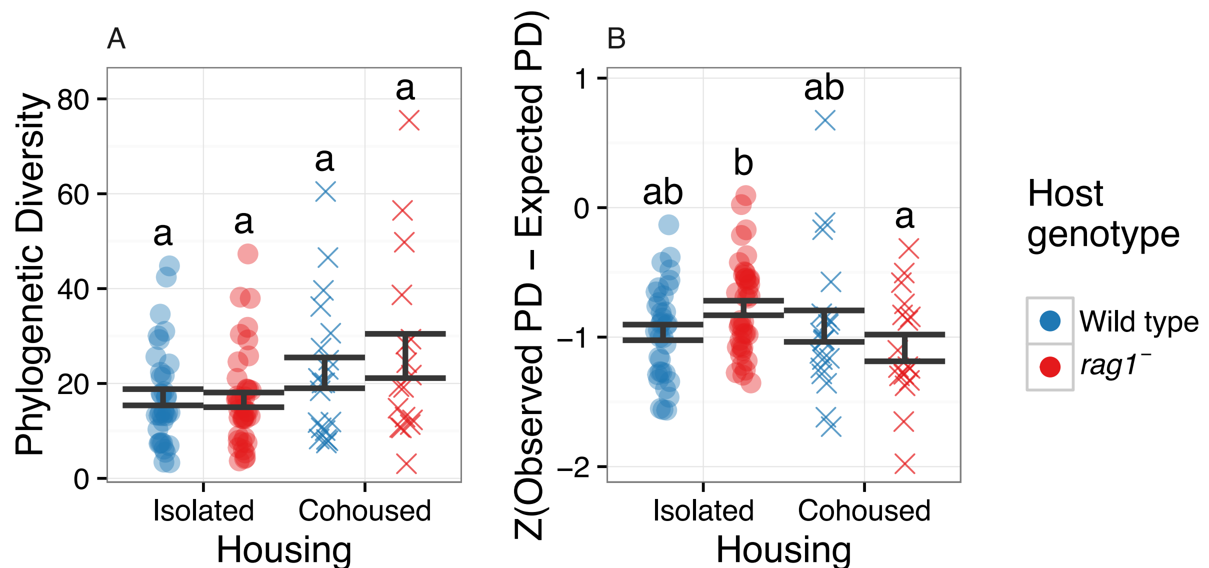


Figure 9. Phylogenetic alpha-diversity of gut communities. **A)** We compared the average phylogenetic diversity (PD) of each experimental treatment, and **B)** the average z-score of the difference between observed and expected PD for each treatment. A z-score below zero means the observed phylogenetic diversity was less than expected from random sampling of the whole community tree given the number of taxa present in the sample. Shared significance letters indicates no significant difference in means

according to the *post hoc* test.

To further investigate the role of adaptive immunity, we used a neutral model to ask whether the presence of adaptive immunity increases the degree of filtering between hosts of each genotype. For each genotype, we asked whether a neutral model could predict the distribution of microbial taxa across hosts by the average abundance of those taxa across hosts of all genotypes and housing treatments. We previously documented a decrease in the fit of a neutral model fit for distributions of zebrafish gut microbiota members through developmental time, including time after the onset of adaptive immunity (Burns *et al.*, 2016). We therefore hypothesized that the presence of adaptive immunity, acting as an ecological filter, is one of the host factors contributing to the reduction of the model fit to the distributions of gut microbial taxa. A lower fit of the model for the wild type compared to *rag1-* hosts in our study would indicate that adaptive immunity increases the strength of filtering between hosts and the source pool of all potential fish associated microbes and thus support our hypothesis.

Our analysis revealed that when genotypes were isolated, the fit of the neutral model was significantly lower for wild type hosts than *rag1*$^-$ hosts, however, when cohoused, we found the opposite relationship (**Figure 10**). In particular, the model fit for wild type hosts was robust to housing (no significant difference), while the fit of the model for *rag1*$^-$ host had a dramatic shift between housing conditions.

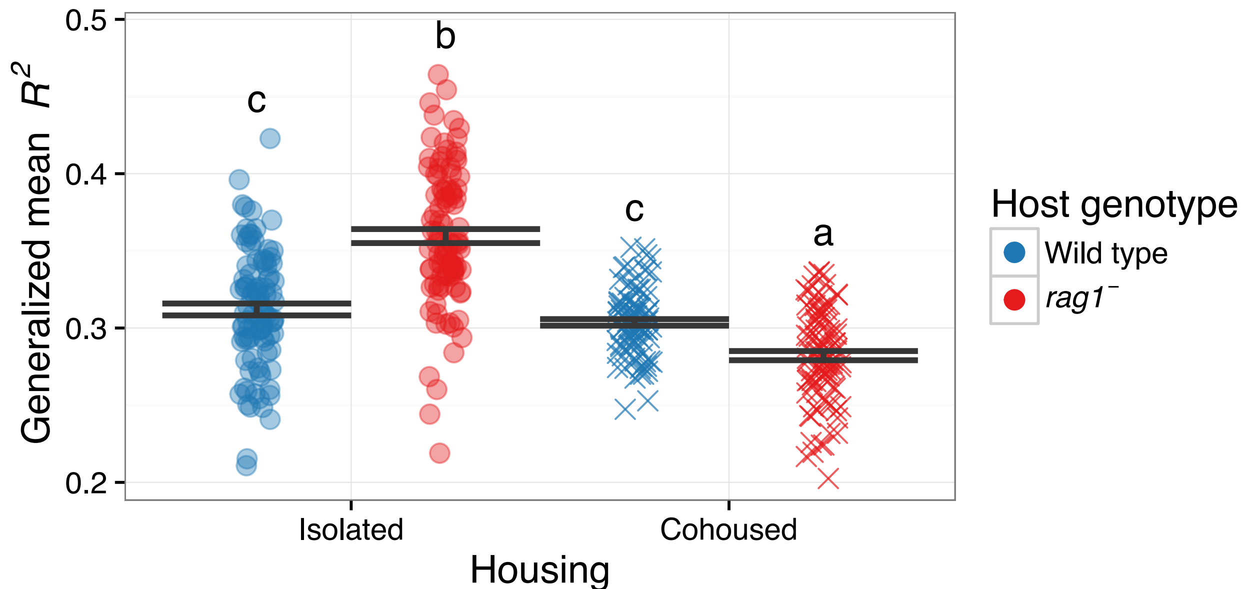


Figure 10. R^2 values for fit of neutral model to each tank. Each point in the plot is one random bootstrap. We then compared the mean R^2 values for the neutral model for each treatment. Shared significance letters indicates no significant difference in means according to the *post hoc* test.

Adaptive immunity acts as an individualizing filter

One of the hallmarks of adaptive immunity is the somatic rearrangement and hypermutation of T and B cell receptors. These processes allow even genetically identical hosts to possess quite different adaptive immune receptor repertoires (Weinstein *et al.*, 2009). Different adaptive repertoires could theoretically filter the gut microbiota in different ways within each wild type host, while other host factors (in a near-isogenic population) should filter the gut microbiota similarly across all hosts. We therefore hypothesized that adaptive immunity acts as an individualizing filter, making wild type hosts more dissimilar, on average, from other wild type hosts than *rag1⁻* hosts would be from each other. When we compared the beta-diversity between hosts of the same genotype, we indeed found that wild type hosts had a significantly greater average dissimilarity than *rag1⁻* hosts (**Figure 11**). For the cohoused genotypes, we expected two possible outcomes. If adaptive immunity is a robust filter, we expected to see no change

in its individualizing effect on the wild type, while if it is not robust, we expected dispersal from *rag1*⁻ hosts to make the wild type hosts more homogenous. If dispersal between hosts was minimal, we expected no change in the homogeneity of either host genotype. As seen in **Figure 11**, when genotypes are cohoused, the homogeneity of wild type hosts remains unchanged from that of isolated wild types, implying that, as an individualizing filter, adaptive immunity is robust to dispersal from other genotypes. On the other hand, *rag1*⁻ hosts become significantly less homogenous (average inter-host similarity goes up) than when they are isolated, a possible mechanism of which is dispersal from wild type hosts causing *rag1*⁻ hosts to become more individualized (**Figure 11**).

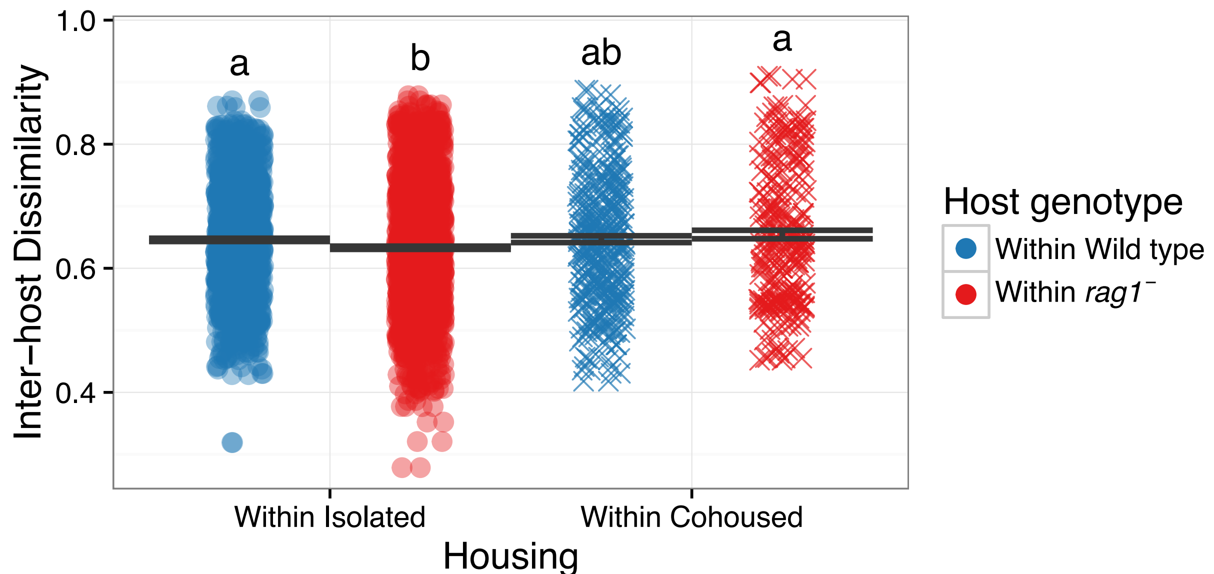


Figure 11: Pairwise Sørensen distances between gut communities. We compared the average pairwise distance between gut communities within experimental treatments. Shared significance letters indicates no significant difference in means according to the *post hoc* test.

Adaptive immunity subtly acts as a different filter from other host factors

Given the above evidence for adaptive immunity as an extra and individualizing filter, we hypothesized that the gut microbiota would be more different on average

between wild type and *rag1⁻* hosts than within each host genotype. However, PERMANOVA analysis of community distance metrics was unable to distinguish gut bacterial communities by genotype, housing condition, or their interaction, and the ordination (**Table 3**), correspondingly, shows no visual separation between groups.

	Df	R^2	p	SES
Housing	1	0.014	0.064	1.62
Host Genotype	1	0.008	0.528	-0.28
Housing:Host Genotype	1	0.009	0.309	0.21
Residuals	110	0.969		
Total	113	1		

Table 3. PERMANOVA results of Sørensen-based pair-wise beta-diversity measures.

We also addressed the hypothesis of adaptive immunity as a non-redundant filter from a different perspective: instead of asking if the presence of adaptive immunity could explain overall differences in gut bacterial community composition, we asked if there were any taxa that were indicative of the presence or absence of adaptive immunity. Using *LefSe* (a linear discriminate analysis that takes taxonomic level assignments into account), we indeed found that in both housing conditions there were taxa that discriminated genotypes. Furthermore, when genotypes were isolated from each other, there were more taxa indicative of each genotype, 33, than when they were cohoused, 14 (**Figure 12**).

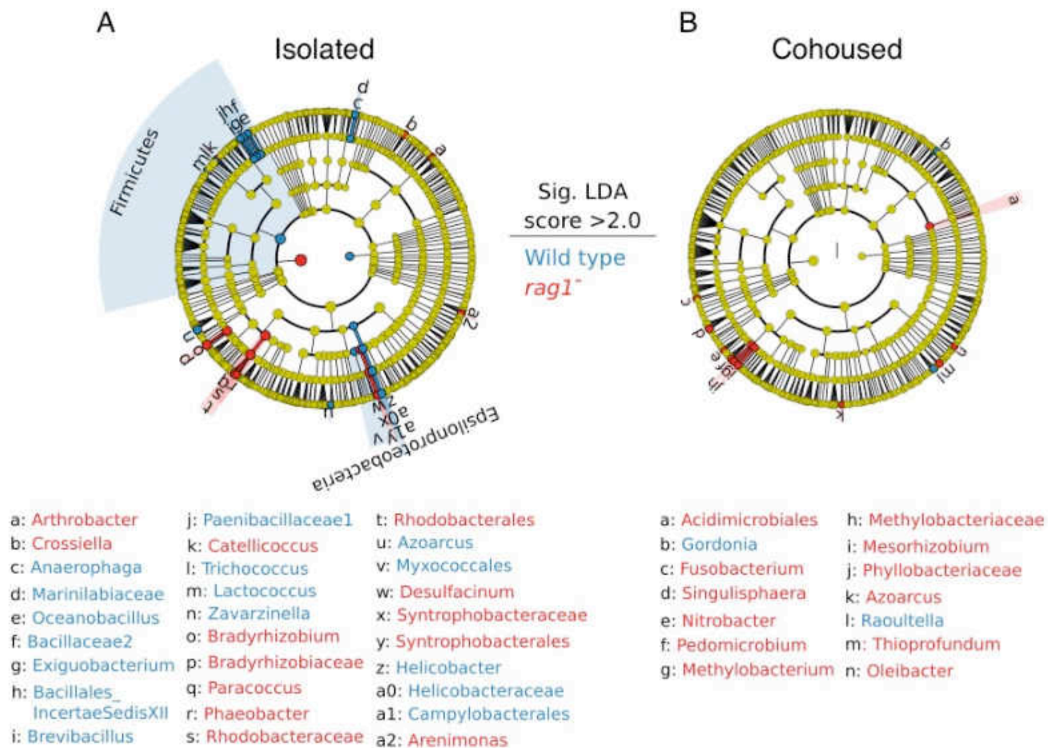


Figure 12. Significant indicator taxa as determined by *LefSe*. Highlighted taxa are overrepresented in the corresponding host genotype by a significant LDA score of >2.0 **A)** for isolated and **B)** cohoused genotypes.

Given our results that supported the hypotheses that adaptive immunity is an extra filter, an individualizing filter, and that there are taxa that discriminate between the presence and absence of adaptive immunity, yet PERMANOVA was unable to distinguish gut bacterial communities by genotype, we returned to the neutral model we used previously in order to determine if there were differences in the subsets of taxa whose distributions deviated from the neutral model. The neutral model predicts that taxon distributions are based solely on differences in dispersal and random sampling of the whole community by each host. It therefore gives us the expectation that taxa that are on average more abundant within each host will be found in more hosts across the population (abundant things are found everywhere). So, in addition to estimating the fit

of the gut microbiota taxa distributions to the neutral model, the analysis can be used to assign each taxon into a "neutral" or "non-neutral" partition based on whether its distribution falls within or outside of, respectively, the 95% confidence interval of the neutral model prediction. The "non-neutral" partition can be further subdivided into "over represented", if the taxon is found in more hosts than expected based on its abundance, and "under represented" if the taxon is found in fewer hosts than expected based on its abundance. Utilizing this model, we divided the gut microbiota of each host into "over represented", "neutral", and "under represented" partitions.

PERMANOVA and ordinations (**Table 4**, **Figure 13**) of the community distances between partitioned communities revealed no main effect of host genotype in the "neutral" partition, and a marginally significant, small effect of genotype within the isolated housing condition. The "neutral" partition comprises the vast majority of the taxa in the gut microbiota, which may explain why PERMANOVA tests on the whole community found no significant differences between genotypes. There were no significant effects of host genotype or housing condition on the "under represented" partition. In the "over represented" partition, however, there was a significant and strong effect of host genotype on the gut microbiota that, unlike all our previous results, was robust to housing condition. That is wild type and *rag1*$^{-}$ genotypes select for different groups of taxa, and this differential selection is maintained even when the potential for microbial and/or host factor transmission between genotypes is increased.

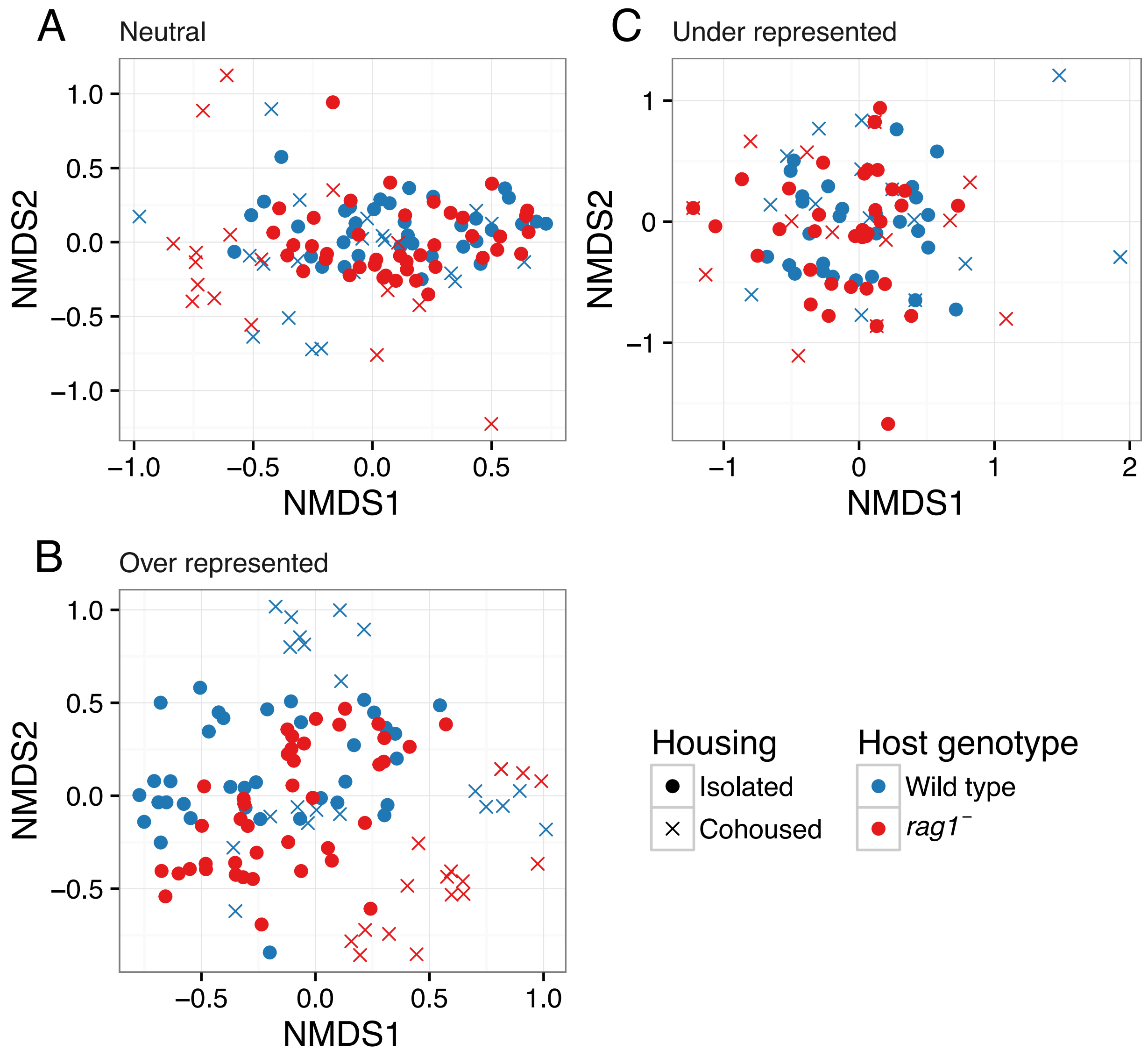


Figure 13. NMDS ordinations of Sørensen distances between gut communities within each neutral model partition. A) Ordination of taxa with distributions within the 95% CI prediction of the neutral model. **B)** Ordination of taxa above the 95% CI prediction of the neutral model. **C)** Ordination of taxa below the 95% CI prediction of the neutral model.

	Df	R^2	p	SES
Partition	2	0.311	0.000	275.43
Housing	1	0.000	1.000	-2.12
Host Genotype	1	0.000	1.000	-2.47
Partition:Housing	2	0.040	0.000	32.62
Partition:Host Genotype	2	0.017	0.000	11.46
Housing:Host Genotype	1	0.001	0.970	-1.34
Partition:Housing:Host Genotype	2	0.015	0.000	10.00
Residuals	316	0.615		
Total	327	1.000		

A

	Df	*R2*	p	SES
Host Genotype in Neutral-Isolated	1	0.005	0.091	1.41
Host Genotype in Neutral-Cohoused	1	0.004	0.145	0.96
Host Genotype in Over represented-Isolated	1	0.009	0.002	5.07
Host Genotype in Over represented-Cohoused	1	0.010	0.001	5.67
Host Genotype in Under represented-Isolated	1	0.004	0.163	0.87
Host Genotype in Under represented-Cohoused	1	0.003	0.524	-0.23
Residuals	321	0.965		
Total	327	1.000		

B

Table 4. PERMANOVA results of ordination factors by Sørensen for all neutral model partitions. A) The full model with partition, housing condition, and host genotype as factors. **B)** Custom contrasts comparing host genotype within housing condition within partition, i.e. a method for a *post hoc* comparison using the PERMANOVA framework.

Discussion

The role adaptive immunity plays in pathogenesis has been well established, but its role in shaping the commensal microbiota is less well understood. Although there is growing evidence that adaptive immunity can act as an ecological filter of the commensal gut bacterial community, the nature and relative strength of this filtering is not known. We endeavored to rigorously test the nature of adaptive immunity's filtering effects utilizing the zebrafish gut microbiota as a model. We observed that adaptive immunity intensifies the filtering between the environmental bacterial community and the host's gut

microbiota. Cohousing wild type and *rag1*$^{-}$ zebrafish completely mitigates this effect such that cohoused wild type hosts become just as similar to the bacterial community in the surrounding water as do *rag1*$^{-}$ hosts. One possible explanation for this phenomenon is that, while adaptive immunity intensifies host filtering of the gut microbiota to some degree, increased transmission from hosts lacking the adaptive immunity filter seeds the water environment with fish-associated microbial taxa. Thus, when both genotypes are cohoused, they appear more similar to the water, because the water is actually more fish-like, even if adaptive immunity is still acting as an environmental filter for the wild type. Another, not mutually exclusive, explanation is that there are differences in host factors (other than adaptive immunity) between wild type and *rag1*$^{-}$ hosts, for example stress hormones such as cortisol, and transmission of these factors reduces the filtering effect of adaptive immunity.

While the above results would seem to suggest that cohousing genotypes causes wild type gut communities to become more like *rag1*$^{-}$ gut communities, the situation may be a bit more complex. As we've shown, while there are no differences in observed phylogenetic diversity between genotypes in either housing condition, cohousing genotypes causes a decrease in the observed phylogenetic diversity relative to expected phylogenetic diversity for *rag1*$^{-}$ gut communities. That is, *rag1*$^{-}$ gut communities are more phylogenetically clustered when cohoused with wild type host than when isolated from them, implying some sort of phylogenetic filtering is occurring in *rag1*$^{-}$ hosts in the presence of immunocompetent hosts. In a similar vein, the neutral model has a better fit to *rag1*$^{-}$ gut community distributions when they are isolated from wild type hosts than when the genotypes are cohoused. The fit of the neutral model to wild type community

distributions, however, is robust to housing condition. Finally, while the data demonstrate that isolated immunocompromised hosts are more similar to each other on average than wild type hosts, this difference disappears in cohoused genotypes. Again, it is not because wild type host become more similar to each other, but rather because *rag1*$^{-}$ hosts become as individualized as wild type hosts.

In addition to determining whether adaptive immunity intensifies host filtering and individualizes hosts, we wanted to know if hosts with functional adaptive immunity filter different taxa than those lacking it. We expected to see this effect manifested as differences in composition of the gut microbiota between genotypes. That is, we expected the average pairwise community distance between genotypes to be greater than within each genotype. However, such analysis on the whole community revealed no significant effects of adaptive immunity. Because we knew there were differences in how well the neutral model fit for each genotype, we used the same model to partition the gut community of each host based on inferred selection. Intriguingly, there was only a marginally significant effect (of small effect size) of host genotype in the "neutral" partition when genotypes were isolated but not cohoused, potentially explaining why there were no measurable effects when considering the whole community. There was, however, a rather large and significant effect of host genotype on the "over represented" partition in both housing conditions, but no significant effect of host genotype at all in the "under represented" partition.

If the primary role of adaptive immunity with regard to the gut microbiota is to filter out more taxa than innate immunity can on its own, we would expect to see stark differences between genotypes in the "under represented" partition. Since the "under

represented" partitions are comprised of those taxa which do not appear consistently across communities, we acknowledge it is possible that the lack of differentiation among underrepresented groups is the result of those groups being too variable to detect. However, we found that there was no significant difference in the dispersion between "over represented", and "under represented" partitions (data not shown), suggesting this explanation is unlikely. Rather, in line with the paradigm popularized in McFall-Ngai (2007) in which the adaptive immune system functions as a facilitator of beneficial constituents of the microbiota, the presence of adaptive immunity changes the composition of microbial taxa that are "over represented", that is, taxa that are present in more hosts than expected from their mean abundance in the source pool. Within the "over represented" partition we see a significant interaction between host genotype and housing condition. As the ordination in **Figure 13B** shows, cohoused genotypes are not a subset of either isolated genotype. This implies some interaction of hosts possessing functional adaptive immunity and those lacking adaptive immunity such that it is not a mere matter of one genotype's community becoming like the other's, but rather a different community assembles in each genotype in the presence of the other. This has important real world implications as most natural populations are comprised of interacting individuals of varying immunocompetence.

Fortunately, we had access to just such a real-world data set with which to explore these ideas. In the next chapter, we used a very similar set of analyses to determine how inflammation, helminth parasite infection, and regional market integration interact to filter the gut microbiota of a group of indigenous people in Ecuador. Similar to the results from our zebrafish, we will show that the filters affecting assembly of the intestinal

microbial community are strongly influenced by the potential for dispersal between hosts, and likely, their environment.

CHAPTER V

MARKET INTEGRATION, CRP LEVELS, AND HELMINTH INFECTION INTERACT TO SHAPE THE GUT MICROBIOTA OF AN INDIGENOUS ECUADORIAN POPULATION

Tara Cepon-Robins collected and extracted DNA for all samples for this chapter. I was responsible for the DNA library preparation and sequencing, as well as the sequence analysis and synthesis of this chapter. James Josh Snodgrass and Melissa Liebert provided input on initial analyses and interpretation. Brendan Bohannan and Karen Guillemin edited this text.

Despite recent strides toward inclusion of a diverse array of subjects (De Filippo *et al.*, 2010; Yatsunenko *et al.*, 2012; Schnorr *et al.*, 2014; Clemente *et al.*, 2015; Martínez *et al.*, 2015), much of what we know about the human gut microbiota comes from subjects primarily in Western, market-integrated, countries. Studies that have focused on differences between market-integrated and more traditionally living populations have found striking differences in their gut microbiotas. In particular, the gut microbial communities of people from market-integrated cultures tend to be less diverse within a person but more dissimilar between people (De Filippo *et al.*, 2010; Yatsunenko *et al.*, 2012; Schnorr *et al.*, 2014; Clemente *et al.*, 2015; Martínez *et al.*, 2015). Some of this reduced within-host (or "alpha") diversity may be explained by the high rates of chronic inflammation found in many market-integrated populations (McDade *et al.*, 2012). Inflammation, both acute and chronic, has been associated with decreased gut microbiota diversity (Spor *et al.*, 2011). Along similar lines, the 'hygiene hypothesis' posits that the lack of diversity of the Western gut microbiota is due in part to decreased

prevalence of helminth infections (Yazdanbakhsh, 2002), which are associated with increased diversity of gut microbial communities (Lee *et al.*, 2014). One of the mechanisms proposed for how helminths increase diversity is through the suppression of the host's inflammatory immune response, but other mechanisms, such as direct competition with or facilitation of bacterial species, have not been ruled out. How immune response, helminth infection, and market integration interact to affect the assembly of the gut microbiota has yet to be elucidated.

The Shuar people of Ecuador provide an intriguing population for studying such an interaction. Shuar living in the Upano Valley (UV) have relatively easy access to market centers and often live in communities comprised of multiple ethnicities. In the region just east of the Upano Valley, separated by the Cutucú Mountains (Cross-Cutucú: CC), the Shuar live, on average, much more traditional lifestyles (Cepon-Robins *et al.*, 2014). These differences in lifestyle are reflected in three style of life (SOL) metrics: House, Market, and Traditional (Liebert *et al.*, 2013). The SOL-House metric is determined by the the coding of amenities in a person's home such as their type of floor, access to running water, or access to electricity. The SOL-Market metric is a percentage of objects owned by an individual from a list of market-associated things, such as a car. Conversely, the SOL-Traditional metric is a percentage of objects owned by an individual from a list of traditional-associated things, like a blowgun. As **Figure 14** shows, these metrics correlate with each region in the way we would expect: SOL-House and SOL-Market are higher in the Upano Valley than the Cross-Cutucú region, and the opposite is true for SOL-Traditional. Both regions experience helminth infection at a relatively high rate (compared to Western populations), while chronic inflammation, as

measured by C-Reactive Protein (CRP), is almost non-existent (McDade *et al.*, 2012). This population, therefore, provides an opportunity for studying how temporary, rather than chronic, differences in CRP levels interact with helminth infection and proximity to market centers to shape the human gut microbial communities. More importantly, though, the Shuar population offers a snapshot of gut microbial communities that have undergone and are undergoing rapid transition (over the course of a few decades) from more traditional selective pressures to more Western ones. We know from anthropological work in the area that, especially for older subjects, their style of life metrics at the point of sample collection are not necessarily the same they would have been earlier in their life. As mentioned above, we also know that there is more opportunity for Shuar in the Upano Valley to have contact, and therefore possible microbial dispersal, with people of various other ethnicities, life styles, and regions of Ecuador than Shuar in Cross-Cutucú, who mostly interact with other Shuar. The data we present below shows that these regional differences can significantly affect the ecological filtering on the gut microbiota by innate immunity and parasitic helminths.

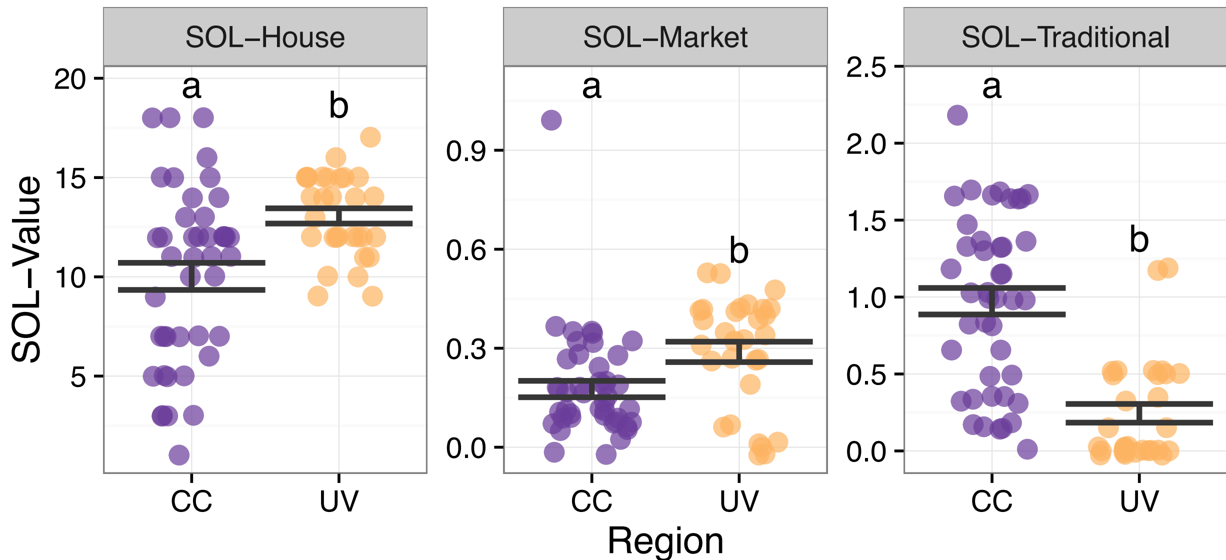


Figure 14. Style of life metrics by region. The SOL-House metric is determined by the the coding of amenities in a person's home. The SOL-Market metric is a percentage of objects owned by an individual from a list of market-associated things. The SOL-Traditional metric is a percentage of objects owned by an individual from a list of traditional-associated things.

Materials and Methods

Stool and blood sampling

The samples for this study were originally collected as a part of the Shuar Health and Life History Project (SHLHP; www.bonesandbehavior.org/shuar), which is co-directed by Larry Sugiyama and Josh Snodgrass. Stool samples were collected as described in (Cepon-Robins *et al.*, 2014). Briefly, subjects were given a pre-packed plastic bag containing an empty stool container and clean implements with which to collect the stool. Subjects returned the containers, and samples were processed within an hour of sample drop off. Stool samples were examined for soil transmitted helminth eggs in the field by a trained observer. Infection intensity levels were classified based on EPG cutoffs according to (Montresor *et al.*, 1998). Blood sampling methods will be fully described in a forthcoming article by Tara Cepon-Robins. In short, several drops of whole blood were collected on filter paper from a finger prick following standard

minimally invasive collection methods (McDade *et al.*, 2007). While in Ecuador, bloodspot cards were stored frozen (at -30 °C) in portable, solar-powered freezers and transported frozen on dry-ice to the US (in order to maximize specimen integrity) for analysis in Dr. Josh Snodgrass' Laboratory at the University of Oregon. Lab analyses were conducted using commercially-available enzyme immunoassays for CRP (M86005M [coating], M86284M [detection]; Biodesign, Memphis, TN), IL-6 (HS600B; R&D Systems, Minneapolis, MN) based on protocols for high-sensitivity assays using dried blood spots (Blackwell *et al.*, 2010, 2011; McDade, 2004; Miller and Mcdade, 2012; Tanner and Mcdade, 2007).

Style of Life

Style of life metrics were determined as described in (Liebert *et al.*, 2013). Briefly, structured interviews, administer mostly in Spanish (or through a bilingual translator for those subjects who did not speak Spanish), were conducted to collect basic demographic and lifestyle information. The selection of items in the Shuar SOL was based on several years of ethnographic observations and pilot testing (unpublished data). The final SOL-Traditional scale contained six items reflecting investment in a foraging lifestyle, while the SOL-Market scale included 12 items reflecting investment in a market economy. Individual scores were calculated as the fraction of list items owned (range 0–1). Six household measures were incorporated as indices of household permanence, access to infrastructure, market participation, and pathogen risk. The SOL-House value for each individual was computed based on a summation of these scores.

DNA Extraction, Library Prep, Sequencing, and Processing

DNA from 300 stool samples was extracted using the QIAamp DNA Stool Mini Kit (Qiagen), following the kit protocol. Library prep of the microbial 16S V4 region was conducted using the same protocol as for the previous chapter. Sequencing was also done on the Illumina (San Diego, CA, USA) HiSeq 2000 platform, generating paired-end 150 nucleotide reads. Sequence quality control, read assembly, and OTU clustering were all done using the same scripts and tool as the previous chapter as well.

Diversity Measures and Statistical Tests

Because interpretation of immune markers, such as CRP, is difficult with regard to children, all analyses were conducted on subjects of 15 years or older. Subjects were only included if they had data for all factors analyzed. Of the 300 samples sequenced, 79 subjects (30 from the Upano Valley; 49 from Cross-Cutucú) met these criteria and were analyzed. All diversity analyses and neutral modeling was done using the same tools and methods in *R* (R Core Team, 2015) as the previous chapter.

Ethics Statement

All participants provided informed consent. Study protocols were approved by village leaders, the Federación Interprovincial de Centros Shuar (FICSH), and the institutional review board of the University of Oregon.

No human genetic data was gathered as part of this project, and the bacterial data gathered was purged of human mitochondrial sequences by removing all sequences matching to the human mitochondrial genome. Genetic material resulting from this study will never be used for commercial cell-line patenting or human DNA research.

Export from Ecuador followed legal guidelines.

Results

There is no significant difference in within-subject diversity between regions

Previous studies that focused on differences between the gut microbiota of market-integrated versus traditional population found that within-subject diversity, i.e. the number and abundance of microbial taxa within an individual host, was reduced in market-integrated populations (De Filippo *et al.*, 2010; Yatsunenko *et al.*, 2012; Schnorr *et al.*, 2014; Clemente *et al.*, 2015; Martínez *et al.*, 2015). Because the Upano Valley has, on average, greater levels of market integration, and lower scores for traditional life styles, we predicted the within-subject diversity to be lower for these individuals than those living in the Cross-Cutucú region (**Figure 15A**). Contrary to these expectations, there was no significant difference in mean phylogenetic diversity (PD) between the two regions.

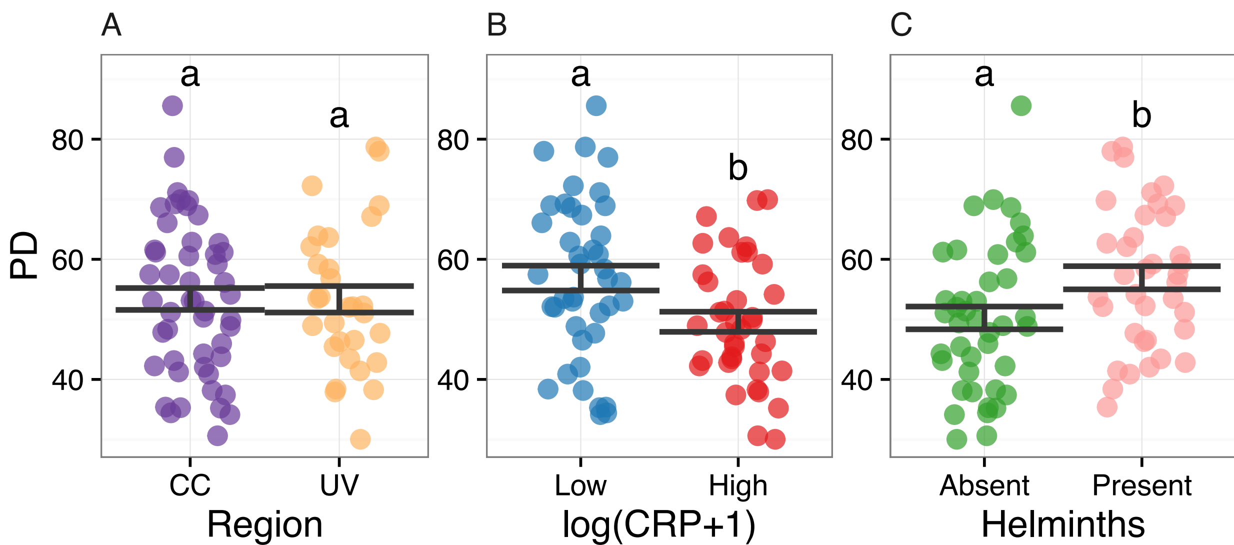


Figure 15. Phylogenetic diversity (PD) by Region, log(CRP+1), and Helminth infection. PD measures the total branch lengths of all taxa within a group, weighted by the abundance of each taxon.

Higher CRP levels are associated with reduced within-subject diversity

CRP is a common marker for inflammation and cardio-vascular disease (McDade *et al.*, 2012). Previous studies have shown that increased inflammation is associated with a reduction in within-subject diversity (Spor *et al.*, 2011), so we predicted that the same trend would be true in our data. As **Figure 15B** shows, Individuals with "high" CRP levels, i.e. greater than the mean CRP score, had significantly lower PD than those with "low" CRP levels, i.e. lower than the mean CRP score.

Helminth infection is associated with increased within-subject diversity

Previous work has found that infection with helminths, particularly *Trichuris* and *Ascaris* species, increase the within-subject diversity of the gut microbiota in humans (Lee *et al.*, 2014). We therefore predicted to see the same results in our study. Indeed, we found that individuals infected with helminths had significantly greater PD, on average, than those with no helminth infection. There were no significant interactions between region, helminth infection or CRP levels with regard to PD (**Table 5**).

	Df	SumSq	MeanSq	Fvalue	Pr(>F)
Region	1	0.05	0.05	0.00	0.985
Helminths	1	978.91	978.91	6.94	0.010
log(CRP+1)	1	715.01	715.01	5.07	0.027
Region:Helminths	1	25.75	25.75	0.18	0.671
Region:log(CRP+1)	1	47.48	47.48	0.34	0.564
Helminths:log(CRP+1)	1	234.94	234.94	1.67	0.201
Region:Helminths:log(CRP+1)	1	7.24	7.24	0.05	0.821
Residuals	71	10016.71	141.08		

Table 5. ANOVA results for PD for region, helminth infection, and CRP levels.

Differences in gut microbiota composition are associated by helminth infection and CRP level more than region

There are two, not mutually exclusive, mechanisms that could lead to differences in gut bacterial community composition between geographic regions: isolation by distance or differences in selective filters. Because of the physical barrier (mountains) between the two regions, which provides the opportunity for isolation by distance, and the difference in market integration, which provides potentially different selective pressures, we expected there to be a difference in gut bacterial community composition between the two regions. However, we found only marginally significant differences between the two regions (**Table 6**). On the other hand, both helminth infection and CRP level were significantly associated with differences in gut microbiota composition, and there was a significant interaction between region and helminth infection.

	Df	R^2	*p*	SES
Region	1	0.03	0.053	1.92
Infection	**1**	**0.05**	**0.007**	**4.18**
log(CRP+1)	**1**	**0.04**	**0.014**	**3.54**
Region:Infection	**1**	**0.03**	**0.032**	**2.39**
Region:log(CRP+1)	1	0.01	0.274	0.23
Infection:log(CRP+1)	1	0.01	0.261	0.26
Region:Infection:log(CRP+1)	1	0.00	0.947	-0.91
Total	78	1.00		
Residuals	71	0.82		

Table 6. PERMANOVA results on Weighted Unifrac distances for factor interactions.

The intensity of Trichuris *infection best explains differences in community composition*

In addition to PERMANOVA (a non-parametric, non-ordination-based method), the *vegan* function *envfit* can be used to fit environmental vectors and factors onto an ordination. **Figure 16** shows the vectors and factors that significantly correlate with community distances according to this function. Similar to the PERMANOVA, helminth infection and CRP levels are significant. The effect of helminth infection, in large part,

seems to be driven by the intensity of *Trichuris*, but not *Ascaris*, infection. This hypothesis is supported, as well, by the results from *LefSe* analysis, a method of linear discriminate analysis (LDA) that takes into account various taxonomic levels. As **Figure 17** shows, whether considering both regions as a whole or separately, most of the significant bio-marker taxa are indicative of either *Trichuris*-only infection or co-infection with both *Ascaris* and *Trichuris* infection. Within just the Upano Valley, most bio-marker taxa are indicative of *Trichuris*-only infection.

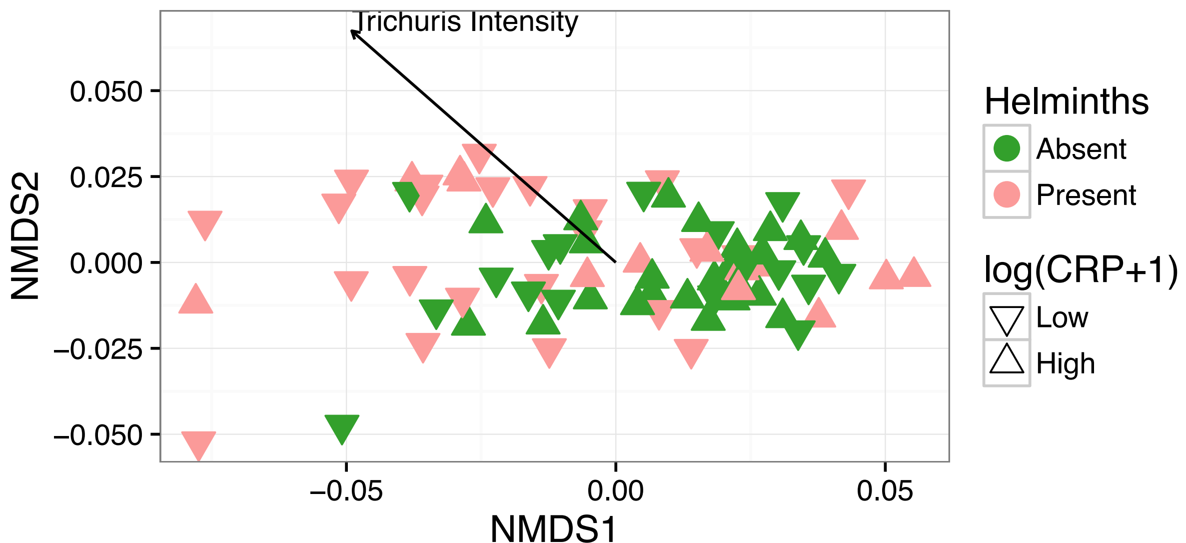
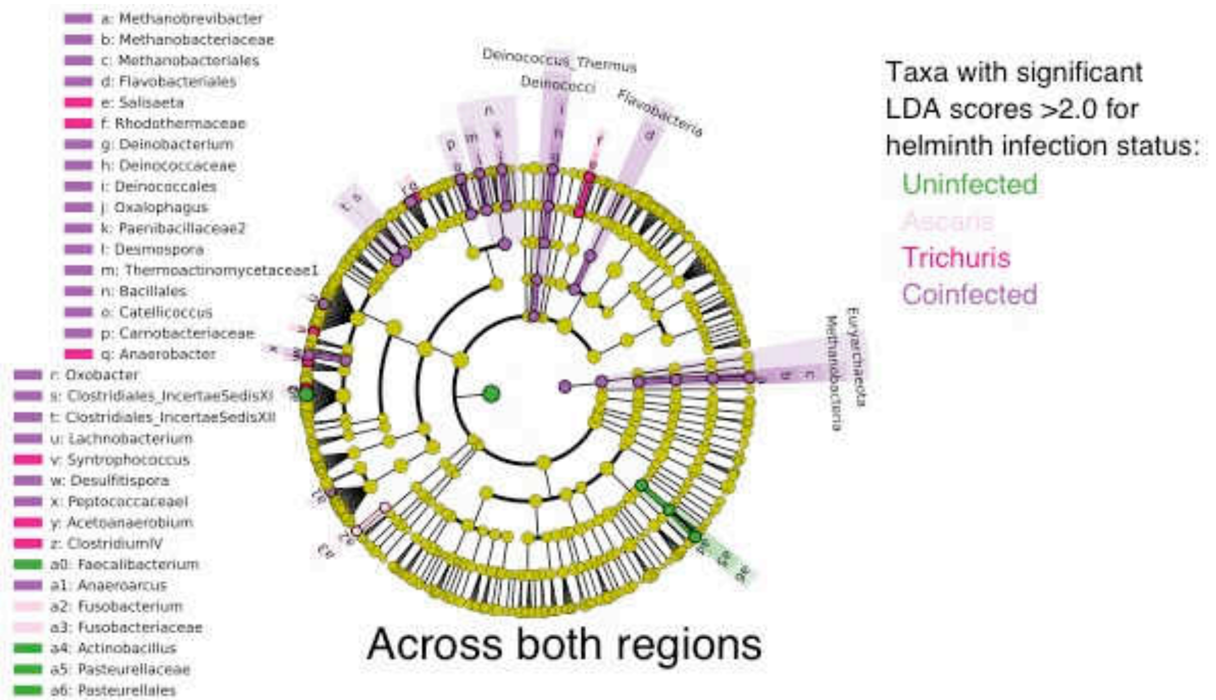


Figure 16. Weighted Unifrac ordinations with significant factors and vectors. A two-dimensional projection of community distances between points. Points closer together are more similar in composition. Point shape and color are determined by the significant factors according the *envfit* function from the *vegan* package in *R*. Significant environmental vectors are shown as black arrows, labelled with the factor name.



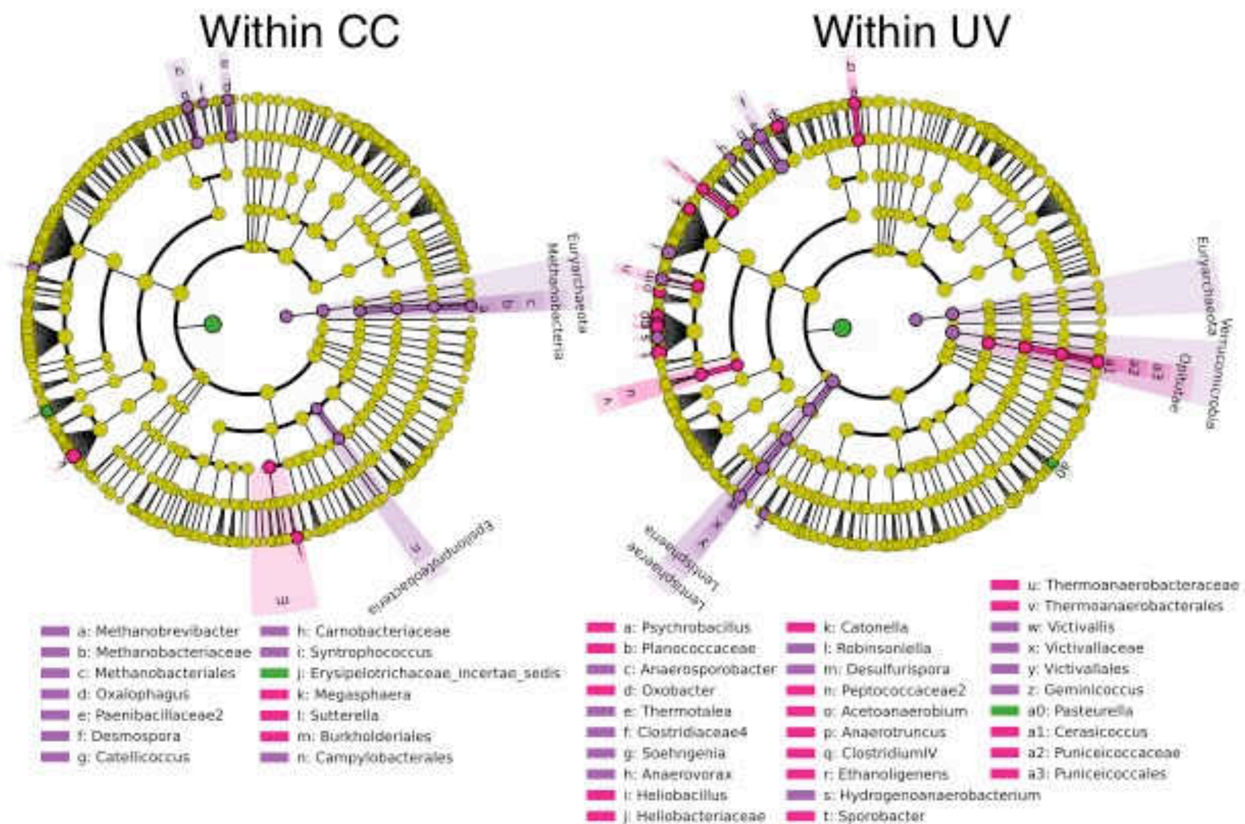


Figure 17: **Indicator taxa as determined by *LefSe* analysis** for various helminth infection statuses across both regions as a whole and within each region on its own.

The Upano Valley is more dissimilar than Cross-Cutucú

A previous study of rural Papua New Guineans (Martínez *et al.*, 2015) found a decrease in average dissimilarity between subjects in that population relative to people from the United States. We therefore hypothesized that, given the greater market integration of the Upano Valley, we should find increased inter-subject dissimilarity in this region relative to Cross-Cutucú. Consistent with our predictions, we found that subjects in the Upano Valley were more dissimilar on average than subjects in Cross-Cutucú (**Figure 18A**).

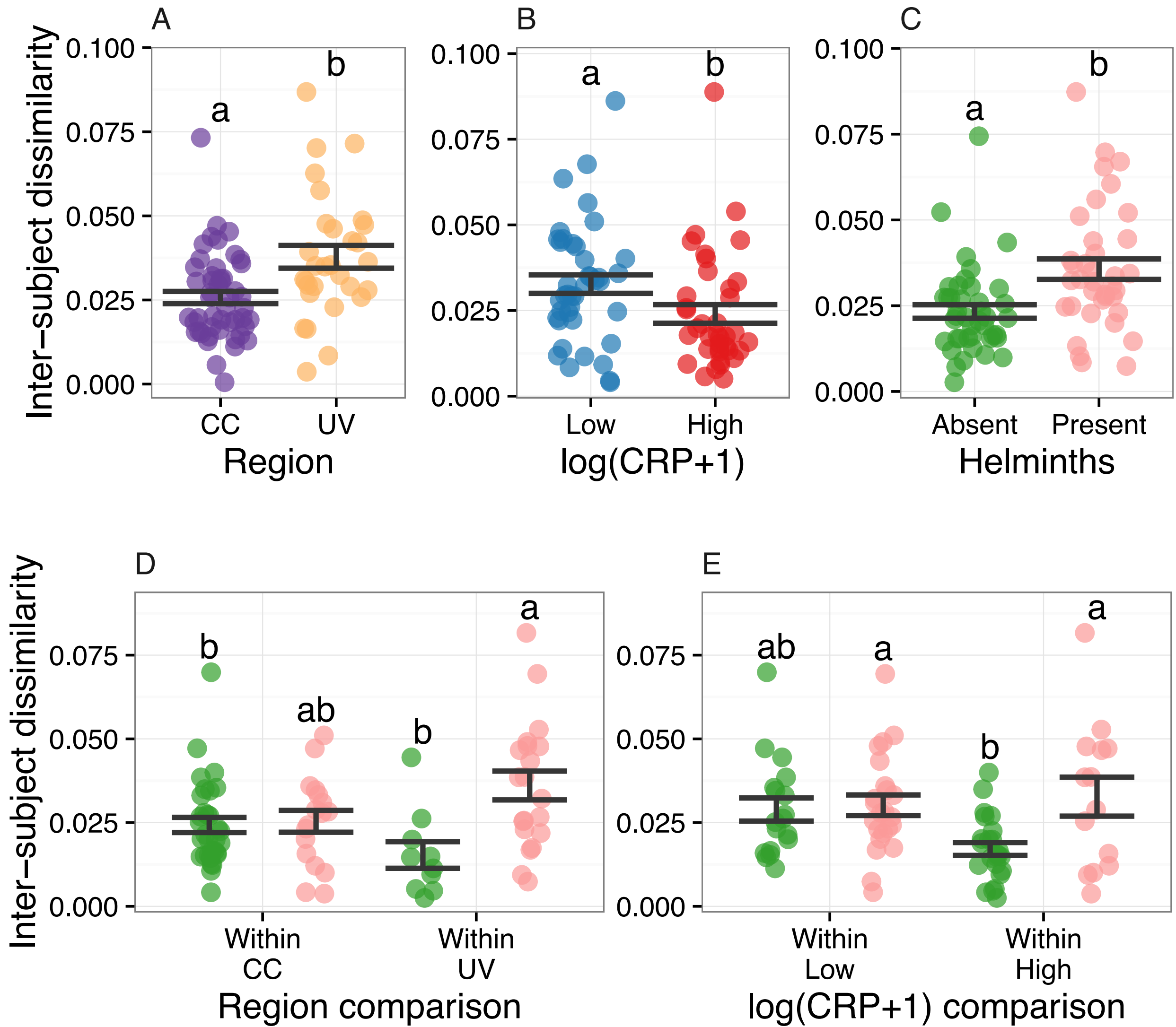


Figure 18. Inter-subject dissimilarity for groups of interest. Inter-subject dissimilarity is determined by the Euclidean distance of each subject to the group centroid using an NMDS-based ordination of weighted Unifrac distances for **A)** region, **B)** log(CRP+1) level, **C)** helminth infection, **E)** the interaction between region and helminth infection, and **F)** the interaction between log(CRP+1) levels and helminth infection.

Higher CRP levels decrease inter-subject dissimilarity

Because CRP is indicative of an inflammatory response, which should manifest similarly across subjects, we expected the selection it imposes on the gut microbiota to be similar across subjects as well. We therefore predicted that people with CRP levels above the mean to have lower inter-subject dissimilarity than people with CRP levels below the

mean. Our analysis revealed that "high" CRP levels were associated with lower average dissimilarity than to "low" CRP levels, supporting our hypothesis (**Figure 18B**).

Helminth infection increases inter-subject dissimilarity in the Upano Valley

We expected that, if helminths act primarily by reducing inflammation, including CRP levels, uninfected individuals should be more homogeneous than infected ones. On the other hand, if helminths act as strong ecological filters, consistently selecting for specific taxa in the gut microbiota, infected individuals should be more homogeneous than uninfected ones. Consistent with our first hypothesis that helminths reduce the effects of inflammation on the gut microbiota, we found that helminth presence increased the average inter-subject dissimilarity (**Figure 18C**). Curiously, this relationship is not consistent across both regions: only in the Upano Valley does helminth infection significantly increase inter-subject dissimilarity, as there is no significant effect of helminth infection in the Cross-Cutucú region (**Figure 18D**). Furthermore, there is no effect of helminth infection on homogeneity when CRP levels are below the mean, but there is when CRP is above the mean. That is to say, when subjects have “low” CRP values, there is no helminth-associated effect on inter-subject dissimilarity, similar to the Cross-Cutucú region. However, subjects with “high” CRP and lacking helminth infection, have significantly lower inter-subject dissimilarity (i.e., they are more homogenous) than any other group. Subjects with “high” CRP levels that are also infected with helminths, however, have levels of inter-subject dissimilarity comparable to people with “low” CRP values.

Neutral processes are more important in Cross-Cutucú than the Upano Valley

Martínez *et al.* (2015) proposed that the reason for the increased inter-subject dissimilarity in the U.S. population compared to the Papua New Guinean population was that increased market integration reduces dispersal between individuals and from the environment to a given individual, allowing differences to arise between individuals due to variable selection and/or ecological drift. Based on this logic, and our similar findings regarding inter-subject dissimilarity (**Figure 18A**), we hypothesized that dispersal between subjects (estimated as the "migration rate" parameter in our neutral assembly model) would be lower in the Upano Valley than Cross-Cutucú. As we found, the average estimated migration rate was lower for the Upano Valley (**Figure 19A**), implying decreased inter-subject dispersal. In addition, we found that the neutral model fit was also lower, implying an increased importance of selection in shaping the gut microbiota of the Upano Valley population versus the Cross-Cutucú population (**Figure 19B**).

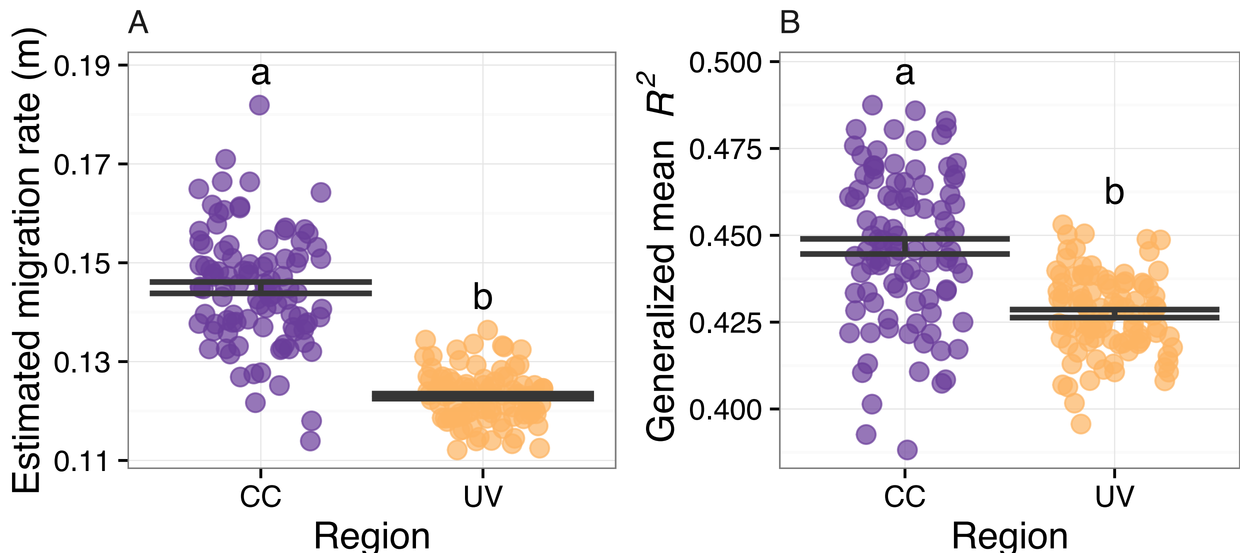


Figure 19. Estimated migration rates and neutral model fit for each region. Each point in the plot represents the neutral model results for a random subsampling, bootstrap, of 27 subjects from each region. The results of the modeling provided **A)** an estimated migration (dispersal) rate for each region and a **B)** model fit, given as the generalized mean R^2.

Discussion

Much of what we know about the effects of inflammation and helminth parasites on the intestinal microbiota are from either Western subjects or mouse studies. Western subjects maintain diets and lifestyles significantly different from other populations across the globe, and laboratory mice are typically maintained in highly hygienic environments that alter their immune development (Beura *et al.*, 2016). Studies that focus on the effect of “westernization” on the gut microbiota often do so by comparing populations that are genetically and geographically distant (De Filippo *et al.*, 2010; Yatsunenko *et al.*, 2012; Schnorr *et al.*, 2014; Clemente *et al.*, 2015; Martínez *et al.*, 2015). The Shuar population of Ecuador offer an intriguing study population because they allow us to compare the effects of immunity, helminth infection, and regional market integration in a set of populations of the same ethnicity that are in geographical proximity.

Contrary to our expectations, we found that phylogenetic within-subject diversity was no different between the market-integrated Upano Valley and the traditional Cross-Cutucú region. This lack of difference could be due to a number of non-mutually exclusive factors. First, it may be that the level of market integration in the Upano Valley is simply not great enough to strongly affect the within-subject diversity of the gut microbiota. Second, it may be that increased market-integration in the Upano Valley is too recent to have had a significant effect on average within-subject diversity. Many of the older population in the Upano Valley most likely experienced lower levels of market-integration earlier in their lives, and if such experiences have legacy effects on the intestinal microbiota, then their microbial within-subject diversity may be higher than expected for their current level of market integration. Third, there may be enough

dispersal between the Upano Valley and Cross-Cutucú to mitigate the differences between the regions.

Despite no difference between regions, our results regarding the effects of CRP levels and helminth infection on PD were consistent with our hypotheses. A probable explanation for the effect of CRP levels on within-subject diversity is that increased inflammation, associated with elevated CRP levels, alters the intestinal environment to favor a small number of taxa that are inflammation tolerant. Regardless of the mechanism, this result is intriguing because the majority of these subjects have CRP levels well below what is considered clinically significant for cardio-vascular disease, and yet these small differences seem to have a measurable effect on the within-subject diversity of the microbiota. The increase in PD due to helminth infection is probably due to similar mechanisms. Helminths are known to secret anti-inflammatory compounds (Zaiss *et al.*, 2015; Ziegler *et al.*, 2015), and as such, their presence may decrease the effects of inflammation, and thus increase the within-subject diversity of the intestinal microbial community. We know from this particular data set that subjects with helminth infections have, on average, lower CRP levels than uninfected subjects (data not shown). Alternatively, helminths could influence the PD of the gut microbiota in an inflammation-independent manner, perhaps by altering the availability of nutrients or other resources.

Along similar lines to the within-subject diversity results, differences in the composition of the gut microbiota, as measured by weighted Unifrac community distances, were primarily driven by helminth infection status and CRP levels, and much less so by region. Again, the lack of differences between regions and significant

differences associated with CRP levels and helminth infection could be due to the similar mechanisms posited above for within-subject diversity. In line with results from other helminth studies (Lee *et al.*, 2014), the intensity of *Trichuris*, rather than *Ascaris*, infection seems to be the main factor associated with compositional differences between subjects, according to ordination-based fitting of factors (**Figure 16**). This result was also born out be the *LefSe* analysis, which found that most of the significant biomarkers were indicative of *Trichuris*-only- or co-infection (**Figure 17**). Consistent with the PERMANOVA analysis, which found a significant interaction with region and infection (**Table 6**), the *LefSe* analysis found fewer biomarkers between helminth infection states in the Cross-Cutucú region than the Upano Valley region. Intriguingly, the archaeal phylum *Euryarcheota*, which contains the H_2-utilizing methanogenic class *Methanobacteria*, was indicative of of helminth co-infection across and within both regions. Increased abundance of these archaea, which are present in a high percentage of human subjects, have been associated with increased energy harvesting by the gut microbiota leading to obesity in humans and mice (Zhang *et al.*, 2009; Samuel and Gordon, 2006). Additionally, many bacterial taxa that indicate various helminth infection states can be found within the phylum *Firmicutes*, also associated with increased energy harvesting and obesity in market-integrated human subjects and mouse hosts. Similarly, when helminth infection was considered as a binary (absent/present), the phylum *Bacteroidetes* was an indicator for helminth absence (data not shown), and this phylum is often associated with the gut microbiota of lean people rather than obese. Given the low rates of obesity found in the Shuar population, these results may require us to reconsider

the relationship between these microbial species, helminth parasites, diet, and overall health outcomes.

Consistent with our hypothesis that increased CRP levels should impose a similar selective pressure on the gut microbiota across subjects, we found that the inter-subject dissimilarity was significantly reduced for people with CRP levels above the mean compared to those with CRP levels below the mean. Likewise, helminth presence was associated with an increase in inter-subject dissimilarity. We have postulated above that a possible mechanism for how helminths alter the intestinal microbial community is through the dampening of the effects of inflammation. Concordantly, we found that for subjects with "low" CRP levels, there was no significant difference in inter-subject dissimilarity whether helminths were present or absent. However, subjects with "high" CRP levels lacking helminth infection were significantly more homogeneous (lower inter-subject dissimilarity) than subjects with "high" CRP levels that had a helminth infection and either group of subjects with "low" CRP levels. While not definitive, these results lend further credence to the hypothesis that helminths alter the gut microbiota, at least in part, by moderating the effects of inflammation.

While there were no or few differences in within-subject diversity or species composition between regions, comparable to other similar studies, we found a significant increase in inter-subject dissimilarity within the Upano Valley relative to the Cross-Cutucú region (**Figure 18A**). This result is similar to results from other studies that found an increase in inter-subject dissimilarity within other market-integrated populations such as the United States or western European countries. As stated above, Martínez *et al.* (2015), proposed that the increase in dissimilarity is due to a decrease in microbial

dispersal between individuals and/or variable selection on the gut microbiota due to differences in lifestyle or genetic factors. When we applied a neutral model of taxonomic distribution, we found that the model fit was lower for the Upano Valley, suggesting that selective processes are more important for community assembly in the Upano Valley than they are in the Cross-Cutucú region (**Figure 19B**). While this is not direct evidence for variable selection, *per se*, it is evidence that there is stronger selection of some kind in the Upano Valley. We also found a reduction in the estimated migration (i.e. dispersal) rate in the Upano Valley (**Figure 19A**), indicating that reduced inter-subject dispersal in market-integrated populations may contribute to increases in inter-subject dissimilarity as well. This difference in dispersal might help explain why there are significant differences in inter-subject dissimilarity between helminth infected and uninfected subjects in the Upano Valley, but not the Cross-Cutucú region. If helminth infection primarily affects the microbiota by reducing the homogenizing selection of inflammation, then when helminths are present, other selective factors or ecological drift might predominate the assembly of the gut microbiota, which may be causing the increased inter-subject dissimilarity seen in helminth-infected people in the Upano Valley. In the Cross-Cutucú region, however, the higher rates of dispersal, which can counter-act ecological drift and variable selection (Vellend, 2010), may mediate these effects such that helminth presence does not have an appreciable effect on inter-subject dissimilarity.

It should be noted that both the decrease in dispersal and neutral processes observed in the Upano Valley Shuar population may be due, in large part, to their increased contact with people of other ethnicities relative to the Cross-Cutucú Shuar population. Dispersal in the Upano Valley Shuar may be lower, not because dispersal

between any individual in the Upano Valley is lower, but rather because the Shuar population is diluted by people of other ethnicities not sampled in this study. Likewise, the increase in the importance of selection in the assembly of the gut microbiota of the Upano Valley Shuar might be due to increased dispersal between the Shuar and other ethnicities who might impart different selection on their microbiota due to differences in lifestyle and genetic background.

In summary, the Shuar populations of the Upano Valley and Cross-Cutucú region provide a unique opportunity to study how the interactions between inflammation, helminth parasite infection, and regional market integration shape the assembly of the intestinal microbial community. In particular, the differences between the market-integration and traditionally living regional populations are not quite as drastic as other studies have found. This may be due to the close geographical proximity and the shared genetic background of each group, or because we are witnessing the effects of market-integration on the gut microbiota as it is occurring, and not after. Market-integration appears to reduce microbial dispersal between individuals and increase the importance of selective factors on the assembly of the gut microbiota. The presence of helminth parasites in the gut community has strong effects on the gut microbiota, and these effects can be amplified by both elevated inflammatory marker levels (CRP), and increased regional market integration. All-in-all, the results of this study provide us with insight into how market-integration changes the relationships among the factors that shape the intestinal microbial community.

CHAPTER VI

CONCLUSION

The interactions between animals and their associated microbial communities are dauntingly complex. Yet, we have begun to appreciate the importance of these interactions in relation to the development, physiology, health, and evolution of all animal species, including our own (Dethlefsen *et al.*, 2007; Clemente *et al.*, 2012; Schluter and Foster, 2012). As I argued in chapter II, this understanding can be enhanced by the application of ecological theory, which was developed specifically to address complex multi-species questions, to the analysis of hosts and their microbiota.

Chapter III provided a specific example of how such an ecological framework—along with the use of a high-throughput, highly controlled experimental model organism—could be used to document the progression of a host-microbe system through developmental time. From these data, as well as from the host-microbe literature, we inferred that host-associated microbial communities are not simply a random sampling of the surrounding microbial environment, but rather that the host plays an active role in filtering which microbes comprise its microbiota.

These data, as well as data from mouse and human studies (Slack *et al.*, 2009; Hooper and Macpherson, 2010) led us to hypothesize that an important filter of the gut microbiota in zebrafish might be the adaptive immune system, a host trait evolved specifically to interact with host-associated microbes. Chapter IV provided the evidence that adaptive immunity is involved in filtering the gut microbiota, but its effects are less pronounced than we had expected. Furthermore, increasing the potential for microbial dispersal between hosts of different genotypes (by cohousing *rag1*$^{-}$ and wild type fish)

altered the nature of host filtering on the microbiota, implying that the nature of a host filter is dependent on the host's interactions with other hosts and the environment.

Working from this basis, chapter V presented a similar analysis applied to a population of indigenous Ecuadorians, the Shuar, living in two regions of differing levels of market-integration, and separated by a significant geographical barrier, the Cutucú mountains. We found that, while within-subject diversity was no different between the regions, there were other hallmarks of "westernization" in the Upano Valley, the region with greater marker-integration. Such hallmarks include increased inter-subject dissimilarity, decreased dispersal of microbes between subjects, and an increase in the relative importance of selective pressures on the microbiota. We also found that even small changes in the inflammation marker had measurable effects on the gut microbiota, and that these effects were mitigated by infection with helminth parasites. Taken together these results may have important implications for human health, especially as more populations across the world become increasingly market-integrated.

For the zebrafish model, the next logical step for future studies would be to determine how both branches of the immune system, the adaptive and innate, act individually and together to filter the gut microbiota. These could be achieved by comparing *myd88*-/- hosts, which have impaired innate immunity, with *rag1*-/- and wild type hosts. Development of a gnotobiotic-adult model for zebrafish could also be used to answer questions regarding the specific functions or taxa being filtered by the host immune system.

With regard to the Shuar study, there is still a large amount of data available for analysis. Examples of such additional data include more health information, such as BMI

and adiposity; differences between sexes, both biological and cultural; and comparisons between children and adults. Taking these factors into account along with the data presented in this dissertation might reveal interesting interactions or stronger microbial filters that could give us an even better understanding of our relationship to our gut microbiota as a species.

APPENDIX

PRIMER SEQUENCES

Oligonucleotide Sequence (1)	**Use**	**Chapter**
CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATXXXXXX**GTGTGCCAGCMGCCGCGG**	16S PCR, round 1	III
ACACTCTTTCCCTACACGACGCTCTTCCGATCTXXXXXX**TACNVGGGTATCTAATCC**	16S PCR, round 1	III
AAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGC	16S PCR, round 2	III
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG	16S PCR, round 2	III
ATGGAGCAATGGCACTGTG	sIgM qPCR	III
CCAAGTCACAAACACCTCCTTGGGC	sIgM qPCR	III
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC	Genome sequencing	III
CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC	Genome sequencing	III
AATGATACGGCGACCACCGAGATCTACACXXXXXXXXTATGGTAATTGT**GTGCCAGCMGCCGCGGTAA**	16S 1-step PCR	IV, V
CAAGCAGAAGACGGCATACGAGATXXXXXXXXAGTCAGTCAGCC**GGACTACNNGGGTNTCTAAT**	16S 1-step PCR	IV, V

1. Bold characters denote the sequences targeting the 16S rRNA gene.
   Underlined Xs illustrate the 6 or 8 base-pair index positions.
   Corresponding colored characters indicate overlapping bases in round 1 and round 2 PCR primers.

REFERENCES CITED

Allan C, Stankey G. (2009). Adaptive environmental management: A practitioner's guide. Springer Dordrecht: Heidelberg.

Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, *et al.* (2011). Enterotypes of the human gut microbiome. *Nature* **473**: 174–80.

Avershina E, Storrø O, Øien T, Johnsen R, Pope P, Rudi K. (2014). Major faecal microbiota shifts in composition and diversity with age in a geographically restricted cohort of mothers and their children. *FEMS Microbiol Ecol* **87**: 280–290.

Bäckhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, *et al.* (2015). Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe* **17**: 690–703.

Bates JM, Mittge E, Kuhlman J, Baden KN, Cheesman SE, Guillemin K. (2006). Distinct signals from the microbiota promote different aspects of zebrafish gut differentiation. *Dev Biol* **297**: 374–386.

Bennet R, Nord CE. (1987). Development of the faecal anaerobic microflora after Caesarean section and treatment with antibiotics in newborn infants. *Infection* **15**: 332–336.

Benson AK, Kelly SA, Legge R, Ma F, Low SJ, Kim J, *et al.* (2010). Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proc Natl Acad Sci* **107**: 18933–18938.

Bergström A, Skov TH, Bahl MI, Roager HM, Christensen LB, Ejlerskov KT, *et al.* (2014). Establishment of intestinal microbiota during early life: A longitudinal, explorative study of a large cohort of Danish infants. *Appl Environ Microbiol* **80**: 2889–2900.

Beura LK, Hamilton SE, Bi K, Schenkel JM, Odumade OA, Casey KA, *et al.* (2016). Normalizing the environment recapitulates adult human immune traits in laboratory mice. *Nature* 1–12.

Blackwell AD, Gurven MD, Sugiyama LS, Madimenos FC, Liebert MA, Martin MA, *et al.* (2011). Evidence for a Peak Shift in a Humoral Response to Helminths: Age Profiles of IgE in the Shuar of Ecuador, the Tsimane of Bolivia, and the U.S. NHANES Yazdanbakhsh M (ed). *PLoS Negl Trop Dis* **5**: e1218.

Blackwell AD, Snodgrass JJ, Madimenos FC, Sugiyama LS. (2010). Life history, immune function, and intestinal helminths: Trade-offs among immunoglobulin E, C-reactive protein, and growth in an Amazonian population. *Am J Hum Biol* **22**: 836–848.

Blankenberg D, Kuster G Von, Coraor N, Ananda G, Lazarus R, Mangan M, *et al.* (2010). Galaxy: A Web-Based Genome Analysis Tool for Experimentalists. In: Vol. Chapter 19. *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc.: Hoboken, NJ, USA, p Unit 19.10.1–21.

Bohnhoff M, Miller CP. (1962). Enhanced Susceptibility to Salmonella Infection in Streptomycin-Treated Mice. *J Infect Dis* **111**: 117–127.

Brotman RM, Ravel J, Cone RA, Zenilman JM. (2010). Rapid fluctuation of the vaginal microbiota measured by Gram stain analysis. *Sex Transm Infect* **86**: 297–302.

Bry L, Falk PG, Midtvedt T, Gordon JI. (1996). A Model of Host-Microbial Interactions in an Open Mammalian Ecosystem. *Science* **273**: 1380–1383.

Burns AR, Stephens WZ, Stagaman K, Wong S, Rawls JF, Guillemin K, *et al.* (2016). Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME J* **10**: 655–664.

Caporaso J, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, *et al.* (2010). QIIME allows analysis of high- throughput community sequencing data. *Nat Methods* **7**: 335–336.

Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, *et al.* (2011). Moving pictures of the human microbiome. *Genome Biol* **12**: R50.

Cepon-Robins TJ, Liebert M a, Gildner TE, Urlacher SS, Colehour AM, Snodgrass JJ, *et al.* (2014). Soil-Transmitted Helminth Prevalence and Infection Intensity Among Geographically and Economically Distinct Shuar Communities in the Ecuadorian Amazon. *J Parasitol* **100**: 598–607.

Chen J, Bittinger K, Charlson ES, Hoffmann C, Lewis J, Wu GD, *et al.* (2012). Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics* **28**: 2106–13.

Clemente JC, Pehrsson EC, Blaser MJ, Sandhu K, Gao Z, Wang B, *et al.* (2015). The microbiome of uncontacted Amerindians. *Sci Adv* **1**: e1500183–e1500183.

Clemente JC, Ursell LK, Parfrey LW, Knight R. (2012). The Impact of the Gut Microbiota on Human Health: An Integrative View. *Cell* **148**: 1258–1270.

Daily GC, Alexander S, Ehrlich PR, Goulder L, Lubchenco J, Matson PA, *et al.* (1997). Ecosystem Services: Benefits Supplied to Human Societies by Natural Ecosystems.

Daniels JA, Lederman HM, Maitra A, Montgomery EA. (2007). Gastrointestinal Tract Pathology in Patients With Common Variable Immunodeficiency (CVID). *Am J Surg Pathol* **31**: 1800–1812.

Decker E, Engelmann G, Findeisen A, Gerner P, Laass M, Ney D, *et al.* (2010). Cesarean Delivery Is Associated With Celiac Disease but Not Inflammatory Bowel Disease in Children. *Pediatrics* **125**: e1433–e1440.

Dethlefsen L, Huse S, Sogin ML, Relman DA. (2008). The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol* **6**: e280.

Dethlefsen L, McFall-Ngai M, Relman DA. (2007). An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* **449**: 811–8.

Dethlefsen L, Relman DA. (2011). Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc Natl Acad Sci* **108 Suppl** : 4554–61.

Dimitriu PA, Boyce G, Samarakoon A, Hartmann M, Johnson P, Mohn WW. (2013). Temporal stability of the mouse gut microbiota in relation to innate and adaptive immunity. *Environ Microbiol Rep* **5**: 200–210.

Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, *et al.* (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci* **107**: 11971–5.

Donskey CJ, Hujer AM, Das SM, Pultz NJ, Bonomo RA, Rice LB. (2003). Use of denaturing gradient gel electrophoresis for analysis of the stool microbiota of hospitalized patients. *J Microbiol Methods* **54**: 249–256.

Duerkop B a, Vaishnava S, Hooper L V. (2009). Immune responses to the microbiota at the intestinal mucosal surface. *Immunity* **31**: 368–376.

Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

Ellis AM, Lounibos LP, Holyoak M. (2006). Evaluating the long-term metacommunity dynamics of tree hole mosquitoes. *Ecology* **87**: 2582–90.

Escherich T. (1988). The Intestinal Bacteria of the Neonate and Breast-Fed Infant. *Clin Infect Dis* **10**: 1220–1225.

Faith DP. (1992). Conservation evaluation and phylogenetic diversity. *Biol Conserv* **61**: 1–10.

Fanning S, Hall LJ, Cronin M, Zomer A, MacSharry J, Goulding D, *et al.* (2012). Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. *Proc Natl Acad Sci* **109**: 2108–2113.

Favier CF, Vaughan EE, De Vos WM, Akkermans ADL. (2002). Molecular Monitoring of Succession of Bacterial Communities in Human Neonates. *Appl Environ Microbiol* **68**: 219–226.

De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, *et al.* (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci* **107**: 14691–14696.

Flajnik MF, Kasahara M. (2010). Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat Rev Genet* **11**: 47–59.

Fox J, Weisberg S. (2011). An {R} Companion to Applied Regression. Second. Sage.

Friswell MK, Gika H, Stratford IJ, Theodoridis G, Telfer B, Wilson ID, *et al.* (2010). Site and Strain-Specific Variation in Gut Microbiota Profiles and Metabolism in Experimental Mice Ahmed N (ed). *PLoS One* **5**: e8584.

Fukami T, Beaumont HJE, Zhang X-X, Rainey PB. (2007). Immigration history controls diversification in experimental adaptive radiation. *Nature* **446**: 436–9.

Fukami T, Nakajima M. (2011). Community assembly: alternative stable states or alternative transient states? *Ecol Lett* **14**: 973–84.

Fukuda S, Toh H, Hase K, Oshima K, Nakanishi Y, Yoshimura K, *et al.* (2011). Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* **469**: 543–7.

Gajer P, Brotman RM, Bai G, Sakamoto J, Schütte UME, Zhong X, *et al.* (2012). Temporal dynamics of the human vaginal microbiota. *Sci Transl Med* **4**: 132ra52.

George LO, Bazzaz FA. (1999). The Fern Understory as an Ecological Filter : Emergence and Establishment of Canopy-Tree Seedlings. *Ecology* **80**: 833–845.

Giardine B. (2005). Galaxy: A platform for interactive large-scale genome analysis. *Genome Res* **15**: 1451–1455.

Goecks J, Nekrutenko A, Taylor J, Galaxy Team T. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* **11**: R86.

Goossens H, Ferech M, Vander Stichele R, Elseviers M. (2005). Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. *Lancet* **365**: 579–87.

Gori A, Tincati C, Rizzardini G, Torti C, Quirino T, Haarman M, *et al.* (2008). Early impairment of gut function and gut flora supporting a role for alteration of gastrointestinal mucosa in human immunodeficiency virus pathogenesis. *J Clin Microbiol* **46**: 757–8.

Van der Gucht K, Cottenie K, Muylaert K, Vloemans N, Cousin S, Declerck S, *et al.* (2007). The power of species sorting: Local factors drive bacterial community composition over a wide range of spatial scales. *Proc Natl Acad Sci* **104**: 20404–20409.

Hartigan J, Hartigan P. (1985). The Dip Test of Unimodality. *Ann Stat* **13**: 70–84.

Hooper L V, Littman DR, Macpherson AJ. (2012). Interactions Between the Microbiota and the Immune System. *Science* **336**: 1268–1273.

Hooper L V, Macpherson AJ. (2010). Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat Rev Immunol* **10**: 159–169.

Hovatter SR, Dejelo C, Case AL, Blackwood CB. (2011). Metacommunity organization of soil microorganisms depends on habitat defined by presence of Lobelia siphilitica plants. *Ecology* **92**: 57–65.

Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, *et al.* (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**: 498–503.

Hubbell SP. (2001). The unified neutral theory of biodiversity and biogeography. In: *Monographs in population biology*. Princeton University Press: Princeton, p xiv, 375 p.

Inman CF, Laycock GM, Mitchard L, Harley R, Warwick J, Burt R, *et al.* (2012). Neonatal colonisation expands a specific intestinal antigen-presenting cell subset prior to CD4 T-cell expansion, without altering T-cell repertoire. *PLoS One* **7**: e33707.

Jakobsson HE, Jernberg C, Andersson AF, Sjölund-Karlsson M, Jansson JK, Engstrand L. (2010). Short-term antibiotic treatment has differing long-term impacts on the human throat and gut microbiome. *PLoS One* **5**: e9836.

Jernberg C, Löfmark S, Edlund C, Jansson JK. (2007). Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *ISME J* **1**: 56–66.

Johansson ME V., Phillipson M, Petersson J, Velcich A, Holm L, Hansson GC. (2008). The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc Natl Acad Sci* **105**: 15064–15069.

Kanther M, Rawls JF. (2010). Host–microbe interactions in the developing zebrafish. *Curr Opin Immunol* **22**: 10–19.

Kawamoto S, Maruya M, Kato LM, Suda W, Atarashi K, Doi Y, *et al.* (2014). Foxp3+ T Cells Regulate Immunoglobulin A Selection and Facilitate Diversification of Bacterial Species Responsible for Immune Homeostasis. *Immunity* **41**: 152–165.

Kelly CP, LaMont JT. (2008). Clostridium difficile--more difficult than ever. *N Engl J Med* **359**: 1932–40.

Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, *et al.* (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**: 1463–1464.

Kerr B, Riley MA, Feldman MW, Bohannan BJM. (2002). Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* **418**: 171–4.

Klepac-Ceraj V, Lemon KP, Martin TR, Allgaier M, Kembel SW, Knapp AA, *et al.* (2010). Relationship between cystic fibrosis respiratory tract bacterial communities and age, genotype, antibiotics and Pseudomonas aeruginosa. *Environ Microbiol* **12**: 1293–1303.

Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, *et al.* (2011). Succession of microbial consortia in the developing infant gut microbiome. *Proc Natl Acad Sci* **108**:

4578–4585.

Koopman J. (2004). Modeling infection transmission. *Annu Rev Public Health* **25**: 303–26.

Kuitunen M, Kukkonen K, Juntunen-Backman K, Korpela R, Poussa T, Tuure T, *et al.* (2009). Probiotics prevent IgE-associated allergy until age 5 years in cesarean-delivered children but not in the total cohort. *J Allergy Clin Immunol* **123**: 335–41.

Lam S., Chua HL, Gong Z, Lam TJ, Sin YM. (2004). Development and maturation of the immune system in zebrafish, Danio rerio: a gene expression profiling, in situ hybridization and immunological study. *Dev Comp Immunol* **28**: 9–28.

Lau MK. (2013). DTK: Dunnett-Tukey-Kramer Pairwise Multiple Comparison Test Adjusted for Unequal Variances and Unequal Sample Sizes. https://cran.r-project.org/package=DTK.

Lederberg J. (2000). Infectious History. *Science* **288**: 287–293.

Lee HH, Molla MN, Cantor CR, Collins JJ. (2010). Bacterial charity work leads to population-wide resistance. *Nature* **467**: 82–5.

Lee SC, Tang MS, Lim YAL, Choy SH, Kurtz ZD, Cox LM, *et al.* (2014). Helminth Colonization Is Associated with Increased Diversity of the Gut Microbiota Davies SJ (ed). *PLoS Negl Trop Dis* **8**: e2880.

Leibold M a., Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, *et al.* (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett* **7**: 601–613.

Lemon KP, Armitage GC, Relman DA, Fischbach MA. (2012). Microbiota-targeted therapies: an ecological perspective. *Sci Transl Med* **4**: 137rv5.

Lichstein JW. (2007). Multiple regression on distance matrices: a multivariate spatial analysis tool. *Plant Ecol* **188**: 117–131.

Liebert M a, Snodgrass JJ, Madimenos FC, Cepon TJ, Blackwell AD, Sugiyama LS. (2013). Implications of market integration for cardiovascular and metabolic health among an indigenous Amazonian Ecuadorian population. *Ann Hum Biol* **40**: 228–242.

Logue JB, Mouquet N, Peter H, Hillebrand H. (2011). Empirical approaches to metacommunities: a review and comparison with theory. *Trends Ecol Evol* **26**: 482–91.

Love MI, Huber W, Anders S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Mackie RI, Sghir A, Gaskins HR. (1999a). Developmental microbial ecology of the neonatal gastrointestinal tract. *Am J Clin Nutr* **69**: 1035S–1045S.

Mackie RIRII, Sghir A, Gaskins HRRR. (1999b). Developmental microbial ecology of the

neonatal gastrointestinal tract. *Am J Clin Nutr* **69**: 1035S–1045.

Macpherson AJ, Geuking MB, Slack E, Hapfelmeier S, McCoy KD. (2012). The habitat, double life, citizenship, and forgetfulness of IgA. *Immunol Rev* **245**: 132–146.

Macpherson AJ, Harris NL. (2004). Opinion: Interactions between commensal intestinal bacteria and the immune system. *Nat Rev Immunol* **4**: 478–485.

Mändar R, Mikelsaar M. (1996). Transmission of Mother's Microflora to the Newborn at Birth. *Neonatology* **69**: 30–35.

Marcobal A, Barboza M, Froehlich JW, Block DE, German JB, Lebrilla CB, *et al.* (2010). Consumption of human milk oligosaccharides by gut-related microbes. *J Agric Food Chem* **58**: 5334–40.

Martínez I, Stegen JC, Maldonado-Gómez MX, Eren AM, Siba PM, Greenhill AR, *et al.* (2015). The Gut Microbiota of Rural Papua New Guineans: Composition, Diversity Patterns, and Ecological Processes. *Cell Rep* **11**: 527–538.

McDade TW. (2004). High-Sensitivity Enzyme Immunoassay for C-Reactive Protein in Dried Blood Spots. *Clin Chem* **50**: 652–654.

McDade TW, Tallman PS, Madimenos FC, Liebert MA, Cepon TJ, Sugiyama LS, *et al.* (2012). Analysis of variability of high sensitivity C-reactive protein in lowland ecuador reveals no evidence of chronic low-grade inflammation. *Am J Hum Biol* **24**: 675–681.

McDade TW, Williams SA (Sharon AA, Snodgrass JJ. (2007). What a Drop Can Do: Dried Blood Spots as a Minimally Invasive Method for Integrating Biomarkers Into Population-Based Research. *Demography* **44**: 899–925.

McFall-Ngai M. (2007). Adaptive Immunity: Care for the community. *Nature* **445**: 153–153.

McFall-Ngai M, Hadfield MG, Bosch TCG, Carey H V, Domazet-Lošo T, Douglas AE, *et al.* (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci* **110**: 3229–3236.

McMurdie PJ, Holmes S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data Watson M (ed). *PLoS One* **8**: e61217.

McMurdie PJ, Holmes S. (2014). Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol* **10**: e1003531.

Mihaljevic JR. (2012). Linking metacommunity theory and symbiont evolutionary ecology. *Trends Ecol Evol* **27**: 323–329.

Miller EM, Mcdade TW. (2012). A highly sensitive immunoassay for interleukin-6 in dried blood spots. *Am J Hum Biol* **24**: 863–865.

Milligan-Myhre K, Charette JR, Phennicie RT, Stephens WZ, Rawls JF, Guillemin K, *et al.* (2011). Study of host-microbe interactions in zebrafish. Third Edit. Elsevier Inc.

Montresor a, Crompton DWT, Hall A, Bundy D a P, Savioli L, Parasitic I. (1998). Guidelines for the evaluation of soil-transmitted helminthiasis and schistosomiasis at community level. *World Heal Organ Rep*.

Moon C, Baldridge MT, Wallace MA, Burnham C-AD, Virgin HW, Stappenbeck TS. (2015). Vertically transmitted faecal IgA levels determine extra-chromosomal phenotypic variation. *Nature* **521**: 90–93.

Murgas Torrazza R, Neu J. (2011). The developing intestinal microbiome and its relationship to health and disease in the neonate. *J Perinatol* **31 Suppl 1**: S29–34.

van Nimwegen FA, Penders J, Stobberingh EE, Postma DS, Koppelman GH, Kerkhof M, *et al.* (2011). Mode and place of delivery, gastrointestinal microbiota, and their influence on asthma and atopy. *J Allergy Clin Immunol* **128**: 948–55.e1–3.

O'Malley MA. (2007). The nineteenth century roots of 'everything is everywhere'. *Nat Rev Microbiol* **5**: 647–51.

Ofiteru ID, Lunn M, Curtis TP, Wells GF, Criddle CS, Francis CA, *et al.* (2010). Combined niche and neutral effects in a microbial wastewater treatment community. *Proc Natl Acad Sci* **107**: 15345–15350.

Oh J, Freeman AF, Park M, Sokolic R, Candotti F, Holland SM, *et al.* (2013). The altered landscape of the human skin microbiome in patients with primary immunodeficiencies. *Genome Res* **23**: 2103–2114.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, *et al.* (2013). vegan: Community Ecology Package.

Oksanen J, Blanchet FGG, Kindt R, Legendre P, Minchin PRR, O'Hara RBB, *et al.* (2016). vegan: Community Ecology Package. https://cran.r-project.org/package=vegan.

Olszak T, An D, Zeissig S, Vera MMP, Richter J, Franke A, *et al.* (2012). Microbial exposure during early life has persistent effects on natural killer T cell function. *Science* **336**: 489–493.

Owens RC, Donskey CJ, Gaynes RP, Loo VG, Muto CA, Owens, Jr. RC, *et al.* (2008). Antimicrobial-Associated Risk Factors for Clostridium difficile Infection. *Clin Infect Dis* **46**: S19–S31.

Paine RT, Tegner MJ, Johnson EA. (1998). Compounded Perturbations Yield Ecological Surprises. *Ecosystems* **1**: 535–545.

Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. (2007). Development of the Human Infant Intestinal Microbiota Ruan Y (ed). *PLoS Biol* **5**: e177.

Pantoja-Feliciano IG, Clemente JC, Costello EK, Perez ME, Blaser MJ, Knight R, *et al.* (2013). Biphasic assembly of the murine intestinal microbiota during early development. *ISME J* **7**: 1112–1115.

Parichy DM, Elizondo MR, Mills MG, Gordon TN, Engeszer RE. (2009). Normal table of postembryonic zebrafish development: Staging by externally visible anatomy of the living fish. *Dev Dyn* **238**: 2975–3015.

Penders J, Thijs C, Vink C, Stelma FF, Snijders B, Kummeling I, *et al.* (2006). Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* **118**: 511–21.

Peterson TS, Ferguson J a, Watral VG, Mutoji KN, Ennis DG, Kent ML. (2013). Paramecium caudatum enhances transmission and infectivity of Mycobacterium marinum and M. chelonae in zebrafish Danio rerio. *Dis Aquat Organ* **106**: 229–39.

Phillips JB, Westerfield M. (2014). Zebrafish models in translational research: tipping the scales toward advancements in human health. *Dis Model Mech* **7**: 739–743.

Prosser JI, Bohannan BJM, Curtis TP, Ellis RJ, Firestone MK, Freckleton RP, *et al.* (2007). The role of ecological theory in microbial ecology. *Nat Rev Microbiol* **5**: 384–92.

R Core Team. (2015). R: A Language and Environment for Statistical Computing. https://www.r-project.org/.

Rauta PR, Nayak B, Das S. (2012). Immune system and immune responses in fish and their role in comparative immunity study: A model for higher organisms. *Immunol Lett* **148**: 23–33.

Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, *et al.* (2011). Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci* **108**: 4680–4687.

Rawls JF, Mahowald MA, Ley RE, Gordon JI. (2006). Reciprocal Gut Microbiota Transplants from Zebrafish and Mice to Germ-free Recipients Reveal Host Habitat Selection. *Cell* **127**: 423–433.

Rawls JF, Samuel BS, Gordon JI. (2004). From The Cover: Gnotobiotic zebrafish reveal evolutionarily conserved responses to the gut microbiota. *Proc Natl Acad Sci* **101**: 4596–4601.

Roberts AP, Mullany P. (2014). Oral biofilms: a reservoir of transferable, bacterial, antimicrobial resistance. *Expert Rev Anti Infect Ther*. http://www.tandfonline.com/doi/abs/10.1586/eri.10.106 (Accessed April 15, 2016).

Roeselers G, Mittge EK, Stephens WZ, Parichy DM, Cavanaugh CM, Guillemin K, *et al.* (2011). Evidence for a core gut microbiota in the zebrafish. *ISME J* **5**: 1595–1608.

Rogers GB, Kozlowska J, Keeble J, Metcalfe K, Fao M, Dowd SE, *et al.* (2014). Functional divergence in gastrointestinal microbiota in physically-separated genetically identical

mice. *Sci Rep* **4**: 5437.

Rogier EW, Frantz AL, Bruno MEC, Wedlund L, Cohen D a, Stromberg AJ, *et al.* (2014). Secretory antibodies in breast milk promote long-term intestinal homeostasis by regulating the gut microbiota and host gene expression. *Proc Natl Acad Sci* **111**: 3074–9.

Salyers AA, Gupta A, Wang Y. (2004). Human intestinal bacteria as reservoirs for antibiotic resistance genes. *Trends Microbiol* **12**: 412–6.

Samuel BS, Gordon JI. (2006). A humanized gnotobiotic mouse model of host-archaeal-bacterial mutualism. *Proc Natl Acad Sci* **103**: 10011–6.

Savage DC. (1977). Microbial ecology of the gastrointestinal tract. *Ann Rev Microbiol* **31**: 107–133.

Savino F, Roana J, Mandras N, Tarasco V, Locatelli E, Tullio V. (2011). Faecal microbiota in breast-fed infants after antibiotic therapy. *Acta Paediatr* **100**: 75–8.

Schaedler RW. (1965). The development of the bacterial flora in the gastrointestinal tract of mice. *J Exp Med* **122**: 59–66.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, *et al.* (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Appl Environ Microbiol* **75**: 7537–7541.

Schluter J, Foster KR. (2012). The Evolution of Mutualism in Gut Microbiota Via Host Epithelial Selection Ellner SP (ed). *PLoS Biol* **10**: e1001424.

Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, *et al.* (2014). Gut microbiome of the Hadza hunter-gatherers. *Nat Commun* **5**: 3654.

Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, *et al.* (2011). Metagenomic biomarker discovery and explanation. *Genome Biol* **12**: R60.

Sekirov I, Finlay BB. (2009). The role of the intestinal microbiota in enteric infection. *J Physiol* **587**: 4159–4167.

Sela DA, Chapman J, Adeuya A, Kim JH, Chen F, Whitehead TR, *et al.* (2008). The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome. *Proc Natl Acad Sci* **105**: 18964–9.

Semova I, Carten JD, Stombaugh J, Mackey LC, Knight R, Farber SA, *et al.* (2012). Microbiota Regulate Intestinal Absorption and Metabolism of Fatty Acids in the Zebrafish. *Cell Host Microbe* **12**: 277–288.

Shea K, Chesson P. (2002). Community ecology theory as a framework for biological invasions. *Trends Ecol Evol* **17**: 170–176.

Shen W, Li W, Hixon JA, Bouladoux N, Belkaid Y, Dzutzev A, *et al.* (2014). Adaptive immunity to murine skin commensals. *Proc Natl Acad Sci* **111**: E2977–E2986.

Slack E, Hapfelmeier S, Stecher B, Velykoredko Y, Stoel M, Lawson MA, *et al.* (2009). Innate and adaptive immunity cooperate flexibly to maintain host-microbiota mutualism. *Science* **325**: 617–620.

Sloan WT, Lunn M, Woodcock S, Head IM, Nee S, Curtis TP. (2006). Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ Microbiol* **8**: 732–740.

Smillie CS, Smith MB, Friedman J, Cordero OX, David LA, Alm EJ. (2011). Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* **480**: 241–4.

Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermudez-Humaran LG, Gratadoux J-J, *et al.* (2008). Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci* **105**: 16731–16736.

Sommer F, Bäckhed F. (2013). The gut microbiota--masters of host development and physiology. *Nat Rev Microbiol* **11**: 227–38.

Sonnenburg JL, Angenent LT, Gordon JI. (2004). Getting a grip on things: how do communities of bacterial symbionts become established in our intestine? *Nat Immunol* **5**: 569–573.

Sørensen T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analysis of the vegetation on Danish commons. *Biol Skr* **5**: 1–34.

Spor A, Koren O, Ley R. (2011). Unravelling the effects of the environment and host genotype on the gut microbiome. *Nat Rev Microbiol* **9**: 279–290.

Stecher B, Robbiani R, Walker AW, Westendorf AM, Barthel M, Kremer M, *et al.* (2007). Salmonella enterica Serovar Typhimurium Exploits Inflammation to Compete with the Intestinal Microbiota Waldor M (ed). *PLoS Biol* **5**: e244.

Stephens WZ, Burns AR, Stagaman K, Wong S, Rawls JF, Guillemin K, *et al.* (2016). The composition of the zebrafish intestinal microbial community varies across development. *ISME J* **10**: 644–654.

Tanner S, Mcdade TW. (2007). Enzyme immunoassay for total immunoglobulin E in dried blood spots. *Am J Hum Biol* **19**: 440–442.

The Human Microbiome Project Consortium. (2012). Structure, function and diversity of the healthy human microbiome. *Nature* **486**: 207–14.

Thiennimitr P, Winter SE, Winter MG, Xavier MN, Tolstikov V, Huseby DL, *et al.* (2011). Intestinal inflammation allows Salmonella to use ethanolamine to compete with the

microbiota. *Proc Natl Acad Sci* **108**: 17480–5.

Thoene-Reineke C, Fischer A, Friese C, Briesemeister D, Göbel UB, Kammertoens T, *et al.* (2014). Composition of Intestinal Microbiota in Immune-Deficient Mice Kept in Three Different Housing Conditions Stover CM (ed). *PLoS One* **9**: e113406.

Trosvik P, Stenseth NC, Rudi K. (2009). Convergent temporal dynamics of the human infant gut microbiota. *ISME J* **4**: 151–158.

Turnbaugh PJ, Bäckhed F, Fulton L, Gordon JI. (2008). Diet-Induced Obesity Is Linked to Marked but Reversible Alterations in the Mouse Distal Gut Microbiome. *Cell Host Microbe* **3**: 213–223.

Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, *et al.* (2009). A core gut microbiome in obese and lean twins. *Nature* **457**: 480–484.

Turnbaugh PJ, Ley RE, Mahowald M a, Magrini V, Mardis ER, Gordon JI. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**: 1027–131.

Uchida D, Yamashita M, Kitano T, Iguchi T. (2002). Oocyte apoptosis during the transition from ovary-like tissue to testes during sex differentiation of juvenile zebrafish. *J Exp Biol* **205**: 711–718.

Urban MC, Leibold MA, Amarasekare P, De Meester L, Gomulkiewicz R, Hochberg ME, *et al.* (2008). The evolutionary ecology of metacommunities. *Trends Ecol Evol* **23**: 311–7.

Vega NM, Allison KR, Khalil AS, Collins JJ. (2012). Signaling-mediated bacterial persister formation. *Nat Chem Biol* **8**: 431–3.

Vellend M. (2010). Conceptual synthesis in community ecology. *Q Rev Biol* **85**: 183–206.

Weinstein JA, Jiang N, White RA, Fisher DS, Quake SR. (2009). High-Throughput Sequencing of the Zebrafish Antibody Repertoire. *Science* **324**: 807–810.

Wickham H. (2009). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wienholds E. (2002). Target-Selected Inactivation of the Zebrafish rag1 Gene. *Science* **297**: 99–102.

Wong S, Stephens WZ, Burns AR, Stagaman K, David LA, Bohannan BJM, *et al.* (2015). Ontogenetic Differences in Dietary Fat Influence Microbiota Assembly in the Zebrafish Gut. *MBio* **6**: e00687–15.

Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA, *et al.* (2011). Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**: 105–8.

Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, *et al.* (2012).

Human gut microbiome viewed across age and geography. *Nature* **457**: 222–227.

Yazdanbakhsh M. (2002). Allergy, Parasites, and the Hygiene Hypothesis. *Science* **296**: 490–494.

Young VB, Schmidt TM. (2004). Antibiotic-Associated Diarrhea Accompanied by Large-Scale Alterations in the Composition of the Fecal Microbiota. *J Clin Microbiol* **42**: 1203–1206.

Yurdusev N, Ladire M, Ducluzeau R, Raibaud P. (1989). Antagonism exerted by an association of a Bacteroides thetaiotaomicron strain and a Fusobacterium necrogenes strain against Clostridium perfringens in gnotobiotic mice and in fecal suspensions incubated in vitro. *Infect Immun* **57**: 724–731.

Zaiss MM, Rapin A, Lebon L, Dubey LK, Mosconi I, Sarter K, *et al.* (2015). The Intestinal Microbiota Contributes to the Ability of Helminths to Modulate Allergic Inflammation. *Immunity* **43**: 998–1010.

Zhang H, DiBaise JK, Zuccolo A, Kudrna D, Braidotti M, Yu Y, *et al.* (2009). Human gut microbiota in obesity and after gastric bypass. *Proc Natl Acad Sci* **106**: 2365–2370.

Zhang H, Sparks JB, Karyala S V, Settlage R, Luo XM. (2015). Host adaptive immunity alters gut microbiota. *ISME J* **9**: 770–781.

Ziegler T, Rausch S, Steinfelder S, Klotz C, Hepworth MR, Kuhl AA, *et al.* (2015). A Novel Regulatory Macrophage Induced by a Helminth Molecule Instructs IL-10 in CD4+ T Cells and Protects against Mucosal Inflammation. *J Immunol* **194**: 1555–1564.

(2010). FASTX Toolkit. http://hannonlab.cshl.edu/fastx_toolkit/index.html.