MICROBIOME COMMUNITY CHANGE IN THE GUTS OF MARINE FISH: FEEDING AND LIFE STAGE TRANSITION AS SIGNIFICANT ORGANIZING FACTORS

A Dissertation

Presented to

The Academic Faculty

By

Darren Joshua Parris

In partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Biology
in the
School of Biological Sciences

Georgia Institute of Technology
December 2018

COPYRIGHT © DARREN JOSHUA PARRIS

MICROBIOME COMMUNITY CHANGE IN THE GUTS OF MARINE FISH: FEEDING AND LIFE STAGE TRANSITION AS SIGNIFICANT ORGANIZING FACTORS

Approved by:

Dr. Frank J Stewart, Advisor
School of Biological Sciences
Georgia Institute of Technology

Dr. Mark Hay
School of Biological Sciences
Georgia Institute of Technology

Dr. Kostas T. Konstantinidis
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Julia Kubanek
School of Biological Sciences
Georgia Institute of Technology

Dr. Tom DiChristina
School of Biological Sciences
Georgia Institute of Technology

Date Approved: November 2, 2018

This work is dedicated to my grandfather, T.W. McCoy Jr., for teaching me to fish

ACKNOWLEDGEMENTS

I would like to thank my wife, Tiffany Parris for her love and support throughout my PhD journey. I would like to thank my parents and grandparents for encouraging me to pursue my passions wherever they may take me. I would also like to thank countless friends and extended family members who have been a part of my life over these last 5 years.

I want to express deep gratitude to my PhD advisor, Dr. Frank Stewart, for his mentorship, friendship, and for being a dedicated scientist I could look to for inspiration. I would also like to thank all the members of the Stewart Lab, past and present, for their support, collaboration, and friendship. I would like to thank countless other scientists, collaborators, instructors, and advisors who have had a major impact on my academic trajectory.

Finally, I would like to thank my PhD committee members, Dr. Kostas Konstantinidis, Dr. Julia Kubanek, Dr. Mark Hay, and Dr. Tom DiChristina for contributing to the progression of this dissertation and to my development as a scientist.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
GI	Gastrointestinal
dpf	Days post fertilization
16S rRNA	Small subunit ribosomal RNA
NGS	Next-generation sequencing
OMZ	Oxygen minimum zone
18S rRNA	Small subunit ribosomal RNA
ETSP	Eastern Tropical South Pacific
C-MORE	Center for Microbial Ecology
R/V	Research Vessel
O2	Oxygen
CTD	Conductivity, temperature, depth probe
PCR	Polymerase chain reaction
NCBI	National Center for Biotechnology
QIIME	Quantitative insights into microbial ecology
OTU	Operational taxonomic unit
PR2	Protist ribosomal reference
PD	Phylogenetic diversity
IACUC	Institutional animal care and use committee
ANOVA	Analysis of variance
SD	Standard deviation
ESV	Exact sequence variant
PCoA	Principal coordinates analysis
SEED	Subsystem database for protein annotation
KEGG	Kyoto Encyclopedia of genes and genomes
BLAST	Basic local alignment search tool
LAB	Lactic acid bacteria
PTS	Phototransferase system

SUMMARY

All animals harbor microbial communities (microbiomes) that play vital roles in host health, development, behavior, and evolution. Determining the processes that regulate microbiome diversity and function is therefore a central question in biology. Numerous investigations have sought to quantify the influence of factors such as diet, host genotype, and environment on gut microbiome assembly, taxonomic composition, and function (Spor et al. 2011, Koenig et al. 2011, Myles et al. 2013). However, these studies have been mostly limited to a handful of model or commercially important host systems. We remain naïve in our understanding of how the importance of different microbiome assembly processes might vary among diverse hosts. This is especially true for the most phylogenetically and ecologically diverse of the vertebrate groups, teleost fishes. In this dissertation, I first describe compositional changes in the gut microbiome associated with the transition from a pelagic larval stage to reef settlement in damselfish (Pomacentridae) and cardinalfish (Apogonidae). Results identify a key transition in microbiome structure across host life stage, suggesting changes in the functional contribution of microbiomes over development in two ecologically dominant reef fish families. Next, I use the clownfish *Premnas biaculeatus* to test how diversity, predicted gene content, and gene transcription of the microbiome vary over a diurnal period following a feeding event. Results confirm feeding as a major restructuring force in intestinal microbiomes over a short timeframe (hours). Finally, I describe ongoing work to characterize the phylogenetic novelty and functional capability of a fish-associated *Endozoicomonas* bacterium. While this genus has been identified as a symbiont of marine invertebrates, its role in the guts of fish remains unknown. Together, these studies

advance our understanding of the diversity and potential function of the fish microbiome, setting the stage for studies to identify the microbiome's effect on fish health and ecology.

CHAPTER 1

INTRODUCTION

All animals harbor microbial communities (microbiomes) that play vital roles in host health, development, behavior, and evolution. Determining the processes that regulate microbiome diversity and function is therefore a central question in biology. Numerous investigations have sought to quantify the influence of factors such as diet, host genotype, and environment on gut microbiome assembly, taxonomic composition, and function (Spor et al. 2011, Koenig et al. 2011, Myles et al. 2013). However, these studies have been limited to a handful of model host systems and often lack the degree of control and replication needed to measure effects of individual structuring processes. We remain naïve in our understanding of how the importance of different microbiome assembly processes might vary among diverse hosts. Knowledge regarding how these factors influence ecology of the gut microbiome from the perspective of microbe-microbe interactions and connectivity between host microbiomes and the surrounding environment is also lacking. This work primarily explores gut microbiome changes in marine fish by describing compositional changes between different developmental stages of reef fish (Chapter 3), quantifying transient effects of feeding events on microbiome composition and function (Chapter 4), and investigating the diversity and functional potential of a mucus/gut specialist Endozoicomonad bacteria (Appendix A).

1.1 Next Generation sequencing to explore microbial communities

The rapid advancement of nucleic acid sequencing over the last 20-30 years has greatly improved our ability to characterize microbial communities from diverse environments by allowing us to bypass the requirement that organisms be isolated in

culture. The development of chain-termination based sequencing in 1977 (Sanger et al. 1977) made it possible for the first time to study microbes by directly sequencing their DNA. Since then, streamlining of the Sanger method and the development of new high-throughput, low cost sequencing techniques like pyrosequencing and Illumina's sequencing by synthesis have made it possible to generate many Gigabytes of nucleic acid data from mixed microbial communities (Metzker 2010). This data can be used for targeted gene studies (e.g. 16S rRNA) to look at taxonomy and phylogeny (Dupperon et al. 2005), genomics/metagenomics to describe functional potential of individual microbes or assemblages (Delong, 2009), and metatranscriptomics to look at gene expression of a given organism or environmental community (Poretsky et al. 2009). Nowhere is the pace of sequencing advancement more apparent than in the field of microbiome research. Numerous molecular studies demonstrate eukaryotes play host to a diverse array of bacteria and archaea that live on exterior and interior body surfaces, especially the GI tract. This "microbiome" putatively serves many functions for the host and we are only beginning to understand host-microbiome ecology for many systems (Turnbaugh et al. 2006; Round & Mazmanian 2009; Heijtz et al. 2011; Greenblum et al. 2012; Theis et al. 2013; Zheng et al. 2013, Wada-Katsumata et al. 2015).

1.2 Microbiome studies in marine fish

The biological significance of host-associated microbiomes is now widely recognized. In vertebrates, the majority of microorganisms inhabit the gastrointestinal (GI) tract at an abundance of potentially trillions of cells whose collective genome may be orders of magnitude larger than that of the host (Whitman et al. 1998; Backhed et al. 2005; Gill et al. 2006). A wealth of studies confirm the importance of a gut microbiome

to host health, fitness, and development (Chung et al. 2012; Lee & Hase 2014), with alterations of microbiome composition affecting such properties as host immunity, susceptibility to toxins, adiposity, efficiency of nutrient acquisition, behavior and mood, and chemical signaling among conspecifics (Turnbaugh et al. 2006; Round & Mazmanian 2009; Heijtz et al. 2011; Greenblum et al. 2012; Theis et al. 2013; Zheng et al. 2013, Wada-Katsumata et al. 2015).

Despite recognition of the role of microbiomes in animal health and ecology, microbiomes, and the processes structuring their assembly, remain unexplored for the vast majority of ecologically relevant taxa. This is true for the most diverse of the vertebrate groups, teleost fishes. Fish account for almost half of all vertebrate species on earth and span a wide spectrum of habitats, host ranges, physiologies, and ecological strategies. Coral reefs, for example, sustain 2500 fish species engaged in a complex web of interactions, including predation, herbivory, corallivory, and symbiosis (Bellwood et al. 2004; Allen 2014). These interactions together modulate material and energy transport and ecosystem structuring on coral reefs, with important consequences for reef-building corals that rely on fish waste for nutrients, or fish herbivory to limit competition with benthic macroalgae (Mumby et al. 2006; Burkepile & Hay 2008; McCauley et al. 2010). Fish have a long evolutionary history and potential co-evolutionary history with symbiotic or commensal microbes. They also have large clutch sizes allowing for high replication among related individuals. All of these traits make fishes ideal systems for microbiome studies.

To date, most studies of fish gut microbiomes have focused primarily on commercial or model species (Roeselers et al. 2011; Clements et al. 2014; Llewellyn et

al. 2014). This research raises the possibility of a core set of gut taxa, notably a dominance by bacteria of the Gammaproteobacteria and Firmicutes, including unique fish-associated strains of the Vibrionales and Clostridia (Spanggaard et al. 2000; Al-Harbi & Naim Uddin 2004; Martin-Antonio et al. 2007; Sullam et al. 2012; Xing et al. 2013; Llewellyn et al. 2014). Significant variation in the fish gut microbiome has also been reported, with changes in taxonomic composition shown to affect host immunity, nutrient acquisition, and epithelial differentiation (Rawls et al. 2004; Bates et al. 2006, 2007; Cheesman & Guillemin 2007; Cheesman et al. 2010; Kanther & Rawls 2010). Fish microbiome composition has been linked to diverse factors including host type (Ye et al. 2014; Givens et al. 2015), trophic ecology and diet (Bolnick et al. 2014a,b; Miyake et al. 2015; Sullam et al. 2015), and environmental conditions (e.g., salinity; Sullam et al. 2012; Schmidt et al. 2015).

1.3 Microbiome assembly

Assembly of a gut microbiome is key to a number of processes during animal development. For example, a single species of Gammaproteobacteria has been shown to promote gut cell proliferation and differentiation in developing zebrafish (Cheesman et al. 2010). Establishment of resident gut microbiota is also important for host immune development and protection from colonizing pathogens (Kanther and Rawls 2010, Albenburg, 2014). Diversity of the early microbiome has been linked to increased likelihood of developing certain metabolic diseases (like Type I diabetes) later in life (Kostic et al. 2015). Therefore, understanding how microbiome assembly progresses within hosts, how diverse factors shape this progression, and how community succession varies over key host life transitions is vital.

Gut microbiome assembly is broadly governed by a balance of stochastic processes where some subset of environmental bacteria randomly colonize the gut, as well as deterministic processes where factors like host genetics, diet, and gut maturation exert selection over the microbes able to persist in the gut. In mammals, early developmental conditions including environment, genetics, and diet (among others) influence microbiome composition in later stages of life (Spor et al. 2011, Koenig et al. 2011, Myles et al. 2013). Colonization of the gut by microbes begins at birth and human infants are seeded with a portion of the vaginal and fecal microbiomes of their mothers during the birthing process (Guiemonde et al. 2006). Inoculation of the microbiome in this way may explain perceived "genetic" differences between individuals whereby siblings tend to have more strains of gut microbes in common than non-siblings (Turnbaugh et al. 2009, Vaishampayan et al. 2010). It could also contribute to stochastic variability in microbiome composition since young are seeded with a random subset of available strains. For obvious reasons, it can be difficult to separate maternal influence from the impact of host genotype on the microbiome. However, higher microbiome similarity between identical rather than fraternal human twins highlights the potential for genotypic effects (Zoetendal et al. 2001, Stewart et al. 2005). Besides the effect of the maternal environment, the rearing environment can lead to differentiation in microbiome composition (Spor et al. 2011). Finally, diet exerts a strong influence on the microbiome and may account for dissimilarity in microbiome composition among children from diverse ethnic backgrounds and between young reared on natural breastmilk vs. formula (Voreades et al. 2015). Despite difficulty disentangling and quantifying the relative

weight of each of these factors on microbiome development, a wealth of mammalian studies have explored this topic.

Although assumed to be similar, much less is known about factors influencing microbiome assembly in fish or key compositional signatures of the fish microbiome at various developmental stages. Fish serve as a good model to study microbiome development since large clutch sizes allow for quantification of inter-individual variation, young develop quickly, and they share many digestive features with mammals (fish still possess a true mouth, esophagus, stomach and intestines). There are also key differences between fish and mammals that make them an interesting subject for microbiome studies. As a group, fish exhibit a much wider range of diet types and levels of diet specialization with the diet of some species consisting of a single food item. Such a high degree of specialization might make some fish species more dependent on their microbiome for certain nutrients as a supplement to their limited diet. Finally, differences in the physical characteristics of aquatic versus terrestrial environments (including microbial load) might lead to fundamentally different patterns of microbiome assembly.

Most fish do not exhibit live birth or parental care and exist in a (comparatively) microbe-enriched environment. Fish first acquire gut microbes directly from the water column and their gut is a simple tube open to the external environment for the first few days of life (Bates et al. 2006). Food may also be an important colonization vector (Nayak 2010) and feeding fish larvae different diets during initial stages of development can influence gut community composition later in life (Ingerslev et al. 2014). Both the environmental and dietary microbial pool fish are exposed to can vary over development since many fish occupy different habitats at certain life stages. For example, most reef

fish have a pelagic larval stage where they spend weeks to month as part of the plankton before settling on reefs. While microbiome composition in adult fish is correlated with a suite of factors including host type (Ye et al. 2014; Givens et al. 2015), trophic ecology and diet (Bolnick et al. 2014a,b; Miyake et al. 2015; Sullam et al. 2015), and environmental conditions (e.g., salinity; Sullam et al. 2012; Schmidt et al. 2015), few studies have addressed how the gut microbiome of fish changes across development. Continued study of fish host systems may provide a better understanding of not only factors controlling microbiome development, but also identify distinct compositional patterns in microbiome colonization including co-occurrence of particular microbial species and succession of microbial communities between life stages.

Stochastic processes, primarily random colonization by environmental microbes, might be expected to dominate fish microbiome assembly at early developmental stages as the gut and immune system mature. As such, it is logical to hypothesize that the prokaryotic community in the gut of larval fishes (or the young of any species) should be more diverse and exhibit greater interindividual variability than in adults. This general pattern is indeed observed in mammal systems (Spor et al. 2011). We might also expect the microbiomes of young fish to share a greater proportion of microbial taxa with communities from the surrounding environment since the resident microbial population is still become established. As the gut and the host mature, selective processes may become important to the persistence of a subset of microbes, decreasing overall species richness and diversity and increasing stability. Alternatively, it could be hypothesized that physical maturation of the gut and greater diet diversity among adult hosts might lead to increased niche partitioning between microbes, promoting increased diversity of gut taxa.

Stephens et al. (2015) conducted one of the only studies addressing this question in fish. In this study, microbiome composition from a single cohort of zebrafish was monitored from 4 days post-fertilization (dpf) to 380 dpf using 16S rRNA gene sequencing. High interindividual variation in microbiome composition was observed among hosts at all stages of development. Distinct compositional shifts were associated with different life stages and diet transitions. Larval fish microbiomes were more similar to the external environment and more diverse than adult hosts, suggesting random colonization promotes diversity in early life. At later stages of development, physical gradients (e.g. oxygen, pH, nutrients) select for only a subset of the gut taxa established early on (Stephens et al. 2015). Similar studies in other fish hosts are needed to determine if these patterns in microbiome diversity throughout developmental are universal or specific to certain host systems.

1.4 Influence of diet on the microbiome

Diet/trophic status is perhaps the most important driver of gut microbiome differentiation, and diet-induced shifts in microbiome composition may influence key aspects of host physiology including adiposity and efficiency of nutrient acquisition (Hooper et al. 2002, Turnbaugh et al. 2006, Cho et al. 2012). There are three primary mechanisms through which diet or trophic status might affect the gut microbiome. First, diet can exert selective pressure by controlling the composition of nutrients making it to the gut (Turnbaugh et al. 2009). Chemical composition of these nutrients may lead to niche partitioning of microbes according to the substrate that is metabolized. Second, feeding ecology can determine the bacterial seed community that animals are exposed to, either via attachment of microbes to incoming food or by influencing where the host

resides in the environment (e.g. benthic vs. pelagic, Grossman 1986, Bouchon-Navaro 1986). Finally, the evolutionary trajectory of hosts from different trophic levels may alter gut physiology. For example, herbivorous fish tend to have longer intestines allowing for increased nutrient absorption (German et al. 2010). Influence of the latter two mechanisms is usually minimized in experimental studies by concentrating on a single host type (Turnbaugh et al. 2008). The vast majority of diet-related microbiome studies focus solely on the impact of defined diet types (e.g. vegetarian vs. non-vegetarian, high-carb vs. low-carb) on prokaryotic community structure, thereby limiting the scope of the research. An emerging question is whether greater taxonomic richness and chemical diversity of dietary components translates to increased taxonomic and functional diversity of microbes in the GI tract. For example, since food-associated microbes are potential colonizers of the gut, we might expect generalist consumers to harbor a larger array of microbial taxa. We might also expect greater chemical diversity in the food of generalists to increase niche availability for assorted microbes. Since richness and diversity of the gut microbiome is positively correlated with host health, understanding processes promoting gut microbial diversity is vital (Ott et al. 2004, Turnbaugh et al. 2009, Larsen et al. 2010).

To date, dietary effects on gut microbiome composition are only known for a handful of fish species. Dietary effects may be particularly important in fish, which as a group exhibit exceptionally high diet diversity due to differences in feeding ecologies, transitions between habitats in some species (e.g., saltwater vs. freshwater), and varying locations for development (benthic vs. pelagic, Lawson et al. 2004, Allen, 2014). Of the datasets attempting to understand the effects of diet on fish gut microbiome composition,

most pool sequence data from multiple studies (all employing slightly different techniques to obtain sequence data) do not include many replicates of the same species to account for individual variation, and categorize fish into discrete feeding categories (Sullam et al. 2012, Givens et al. 2015). A study by Bolnick et al. (2014) asked whether diet diversity in sticklebacks was positively related to microbial diversity in the gut. In this study, 16S rRNA gene sequencing was used to examine microbiome composition from field-collected fish and experimental feeding treatments. Diet diversity in the field was quantified indirectly using stable isotope signatures to differentiate individual stickleback into pelagic vs. benthic microhabitats and to identify fish spending various amounts of time in both habitats. Importantly, this metric of diet diversity assumed that individuals spending an equal amount of time in pelagic vs. benthic environments exhibited greater diet diversity. In the lab, fish were fed either chironomid larvae or daphnia, or an equal mixture of the two as standardized diet diversity treatments. These treatments are highly restricted when compared to expectations of natural diet diversity. Interestingly, results from the field and laboratory experiments showed a negative relationship between diet diversity and gut microbial diversity leading the authors to conclude that indirect effects of diet can interact non-additively to affect prokaryotic diversity in the GI tract. For example, some food sources might contain chemicals inhibitory to certain taxa.

An overlooked aspect of the relationship between diet and microbiome structure are feeding effects independent of diet type. For instance, what are the coupled metabolic interactions that define a food processing event (digestion) and how does feeding frequency affect microbiome structure? It is easy to conceptualize how physical

gradients like oxygen concentration structure communities in the intestine, as oxygen decreases sharply as you move from the submucosal surface of the intestine to the center of the lumen (Espey, 2013). This creates microhabitats with aerobic taxa lining the walls of the intestine and anaerobic or facultative anaerobic taxa in the mid-lumen (Figure 3, Espey 2013). Similar to what is shown in chemostat cultures, local consumption of oxygen by aerobic taxa following a feeding event might promote conditions favorable for anaerobic metabolism creating temporal niches for other microbes (Gerritse et al. 1990). In addition, specific enzymes and metabolites produced by one microbe during digestion might be used as an energy source for another (Comstock and Coyne, 2003). Temporally organized metabolic coupling has been documented in other ecosystems. In planktonic marine environments, for example, heterotrophic bacteria have been shown to time their metabolism around activity of phototrophic community members (Otteson, 2014). These types of temporal organization of gene expression (i.e. timing metabolism around different stages of the digestive process) may help explain the enormous diversity of gut microbes.

The frequency of feeding might also affect the magnitude of microbiome change in response to feeding. Continuous or near-continuous feeding, by grazing animals for example, may maintain relatively constant substrate conditions in the gut as well as a steady stream of food-associated microbes, and therefore a stable assemblage of microbes with slight compositional shifts post-feeding. In contrast, intermittent feeding may promote large compositional changes associated with the transition from a relatively inactive, but stable, "fasting" microbiome to a "bloom" community after feeding. In vertebrate guts, fasting/feeding cycles have been shown to drastically alter the abundance

of individual bacterial groups, with fasting associated with higher occurrence of Bacteriodetes (Crawford et al. 2009, Kohl et al. 2014). In addition to altering the magnitude of microbiome change in response to feeding, differences in feeding regime have been shown to affect baseline microbiome composition and metabolite production (Thais et al. 2014, Li et al. 2017). However, the factors driving microbiome differences linked to feeding regime are likely complex, and potentially related to changes in host physiology (Secor and Carey 2016), as well as sampling the microbiome at varying stages in the digestive cycle.

1.5 Objectives

Microbiomes remain vastly underexplored for a majority of ecologically relevant taxa. This is true for the most diverse group of vertebrates, teleost fish. Outside of a handful of model or commercially important species, we know very little about the microbiome of marine fish. The primary objective of this work is to characterize and explore determinants of gut microbiome structure in reef fish (Chapters 3, 4, and Appendix A). Chapter 2 describes another poorly understood system, microeukaryotes in low oxygen zones of the Ocean. The work in chapter 2 allowed me to look at the effect of environmental factors (notably oxygen) on microbiome structure rather than host association. Chapter 2 also represents my first endeavor using NGS techniques to characterize microbial communities and many of the statistical analyses used for studying environmental communities are conducted similarly in later chapters.

1.5.1 Objective 1 (Chapter 2)

Marine oxygen minimum zones (OMZs) support complex microbial communities adapted for life under low oxygen conditions. While bacteria and archaea in OMZs are becoming better understood, and numerous studies have examined benthic meiofaunal diversity beneath OMZs (Neira et al. 2001, Levin 2003), little is known about pelagic eukaryote diversity in this unique oxygen-depleted environment. Here, we use 18S rRNA gene amplicon sequencing to provide a snapshot of eukaryote community structure in two size fractions (0.2-1.6 µm, >1.6 µm) along a depth gradient through the ETSP OMZ off northern Chile. To our knowledge, this is the first community-level genetic survey of microbial eukaryotes in this anoxic water column. While the topic of this Chapter is distinct from the rest of this dissertation, it represents my first primary author work using NGS techniques to describe microbial communities. The methodological and statistical techniques learned and applied here are utilized in subsequent chapters which are the focus of this work.

1.5.2 Objective 2 (Chapter 3)

The Pomacentridae (damselfish) and Apogonidae (cardinalfish) are among the most common fish families on coral reefs and in the aquarium trade. Members of both families undergo a pelagic larvae phase prior to settlement on the reef, where adults play key roles in benthic habitat structuring and trophic interactions. Fish-associated microbial communities (microbiomes) significantly influence fish health and ecology, yet little is known of how microbiomes change with life stage. Here, our objective was to quantify the taxonomic (16S rRNA gene) composition of whole gut microbiomes from ten species of damselfish and two species of cardinalfish from Lizard Island, Australia,

focusing specifically on comparisons between pelagic larvae prior to settlement on the reef versus post-settlement juvenile and adult individuals. Results identify a key transition in microbiome structure across host life stage, suggesting changes in the functional contribution of microbiomes over development in two ecologically dominant reef fish families.

1.5.3 Objective 3 (Chapter 4)

Diet is a major determinant of intestinal microbiome composition. While studies have evaluated microbiome responses to diet variation, less is understood of how the act of feeding influences the microbiome independent of diet type. Here, we use the clownfish *Premnas biaculeatus*, a species reared commonly in ornamental marine aquaculture, to test how the diversity, predicted gene content, and gene transcription of the microbiome vary over a two-day diurnal period with a single daily feeding event. This study used fish fed four times daily, once daily, or every three days prior to the diurnal period, allowing us also to test how feeding frequency affected microbiome diversity. Results showed little effect of long-term feeding frequency but confirm feeding as a major restructuring force in intestinal microbiomes over a short timeframe (hours).

1.5.4 Objective 5 (Appendix A)

Gammaproteobacteria of the Endozoicomonas genus are commonly detected in microbiome studies of invertebrates and we show in Chapter 3 that they are persistent, core members of the gut microbiome of marine fish. In invertebrates including corals, sponges, and sea slugs Endozoicomonads have been described as putative symbionts with functions ranging from carbohydrate fermentation to antibiotic production. My objective

here was to characterize the genome of a potentially novel, fish-associated Endozoicomonad isolated by single-cell sorting from gut contents of the oval butterflyfish. We provide phylogenetic evidence that this is a unique species of Endoizoicomas bacterium and compare functional potential of this genome to available Endozoicomonas genomes isolated from marine invertebrates. This information sets the stage for exploration of Endozoicomonas in other marine vertebrates and serves as a starting point for comparative genomic studies on vertebrate-associated Endozoicomonas genotypes.

1.6 References

Adams, S. M., Lindmeier, J. B., & Duvernell, D. D. (2006). Microsatellite analysis of the phylogeography, Pleistocene history and secondary contact hypotheses for the killifish, Fundulus heteroclitus. Molecular ecology, 15(4), 1109-1123.

Albenberg, L. G., & Wu, G. D. (2014). Diet and the intestinal microbiome: associations, functions, and implications for health and disease. Gastroenterology, 146(6), 1564-1572.

Al-Harbi, A. H., and M. Naim Uddin. (2004). Seasonal variation in the intestinal bacterial flora of hybrid tilapia (Oreochromis niloticus×Oreochromis aureus) cultured in earthen ponds in Saudi Arabia. Aquaculture, 229: pp. 37-44.

Allen, G. R. (2014). Review of Indo-Pacific coral reef fish systematics: 1980 to 2014. Ichthyological Research, 62(1), pp. 2-8.

Backhed, F., Ley, R., Sonnenburg, J., Peterson, D., and Gordon, J. (2005). Host-Bacterial Mutualism in the Human Intestine. Science, 307, pp. 1915-1920.

Bates, J.M., Mittge, E., Kuhlman, J., Baden, K. N., Cheesman, S. E., & Guillemin, K. (2006). Distinct signals from the microbiota promote different aspects of zebrafish gut differentiation. Dev Biol 297, pp. 374–386.

Bates J. M., Akerlund, J., Mittge, E., & Guillemin, K. (2007). Intestinal alkaline phosphatase detoxifies lipopolysaccharide and prevents inflammation in zebrafish in response to the gut microbiota. Cell Host Microbe 2, pp. 371–382.

Bellwood DR, Hughes TP, Folke C, & Nyström M. (2004). Confronting the coral reef crisis. *Nature*. 429, pp. 827- 833.

Boas, M., Feldt-Rasmussen, U., Skakkebæk, N. E., & Main, K. M. (2006). Environmental chemicals and thyroid function. European Journal of Endocrinology, 154(5), 599-611.

Bolnick DI, Snowberg LK, Hirsch PE, Lauber CL, Knight R, Caporaso JG, Svanbäck R. (2014a). Individuals' diet diversity influences gut microbial diversity in two freshwater fish (threespine stickleback and Eurasian perch. Ecol Lett., 17(8), pp. 979-87.

Bolnick, D. I., Snowberg, L. K., Hirsch, P. E., Lauber, C. L., Parks, B., Lusis, A. J., ... & Svanbäck, R. (2014b). Individual diet has sex-dependent effects on vertebrate gut microbiota. Nature Communications, 5.

Bouchon-Navaro, Y. (1986). Partitioning of food and space resources by chaetodontid fishes on coral reefs. Journal of Experimental Marine Biology and Ecology, 103(1), 21-40.

Burkepile DE, & Hay ME. (2008). Herbivore species richness and feeding complementarity affect community structure and function on a coral reef. *Proceedings of the National Academy of Sciences*, 105, p. 16201.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., ... & Huttley, G. A. (2010). QIIME allows analysis of high-throughput community sequencing data. Nature methods, 7(5), 335-336.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., ... & Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences, 108(Supplement 1), 4516-4522.

Cheesman S. and Guillemin K. (2007). We know you are in there: conversing with the indigenous gut microbiota. Res Microbiol 158, pp. 2–9.
Cheesman S. E., Neal, J. T., Mittge, E., Seredick, B. M., & Guillemin, K. (2010). Microbes and Health Sackler Coloquiium: Epethelial cell proliferation in the developing zebrafish intestine is regulated by the Wnt pathway and microbial signaling via Myd88. Proc Natl Acad Sci. 108, pp. 4570–4577.
Cho, I., & Blaser, M. J. (2012). The human microbiome: at the interface of health and disease. Nature Reviews Genetics, 13(4), 260-270.
Chung H, Pamp SJ, Hill JA, Surana NK, Edelman SM, Troy EB, Reading NC, Villablanca EJ, Wang S, Mora JR, Umesaki Y, Mathis D, Benoist C, Relman DA, Kasper DL. (2012). Gut immune maturation depends on colonization with a host-specific microbiota. Cell, 149(7), pp. 1578-93.
Clements KD, Angert ER, Montgomery WL, Choat JH. (2014). Intestinal microbiota in fishes: what's known and what's not. Molecular Ecology. 23:1891-1898.
Comstock, L. E., & Coyne, M. J. (2003). Bacteroides thetaiotaomicron: a dynamic, niche‐adapted human symbiont. Bioessays, 25(10), 926-929.
Corse, E., Costedoat, C., Chappaz, R., Pech, N., MARTIN, J. F., & Gilles, A. (2010). A PCR‐based method for diet analysis in freshwater organisms using 18S rDNA barcoding on faeces. Molecular Ecology Resources, 10(1), 96-108.
DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., ... & Andersen, G. L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Applied and environmental microbiology, 72(7), 5069-5072.
Duperron, S., Nadalig, T., Caprais, J. C., Sibuet, M., Fiala-Médioni, A., Amann, R., & Dubilier, N. (2005). Dual symbiosis in a Bathymodiolus sp. mussel from a methane seep on the Gabon continental margin (Southeast Atlantic): 16S rRNA phylogeny and distribution of the symbionts in gills. Applied and environmental microbiology, 71(4), 1694-1700.
Espey, M. G. (2013). Role of oxygen gradients in shaping redox relationships between the human intestine and its microbiota. Free Radical Biology and Medicine, 55, 130-140.
Furukawa, K., & Fujihara, H. (2008). Microbial degradation of polychlorinated biphenyls: biochemical and molecular features. Journal of bioscience and bioengineering, 105(5), 433-449.
Ganesh, S., Parris, D. J., DeLong, E. F., & Stewart, F. J. (2014). Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. The ISME journal, 8(1), 187-211.
German, D. P., Nagle, B. C., Villeda, J. M., Ruiz, A. M., Thomson, A. W., Balderas, S. C., & Evans, D. H. (2010). Evolution of herbivory in a carnivorous clade of minnows (Teleostei: Cyprinidae): effects on gut size and digestive physiology. Physiological and Biochemical Zoology, 83(1), 1-18.
Gerritse, J., Schut, F., & Gottschal, J. C. (1990). Mixed chemostat cultures of obligately aerobic and fermentative or methanogenic bacteria grown under oxygen-limiting conditions. FEMS Microbiology Letters, 66(1-3), 87-93.
Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, Gordon JI, Relman DA, Fraser-Liggett CM, Nelson KE (2006). Metagenomic analysis of the human

distal gut microbiome. Science 312, pp. 1355–1359.
Givens, C. E., Ransom, B., Bano, N., & Hollibaugh, J. T. (2015). Comparison of the gut microbiomes of 12 bony fish and 3 shark species. Marine Ecology Progress Series, 518, pp. 209-223.
Greenblum, S., Turnbaugh, P., and Borenstein, E. (2012). Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. Proc. Natl. Acad. Sci. 109, pp. 594–599.
Grossman, G. D. (1986). Food resource partitioning in a rocky intertidal fish assemblage. Journal of Zoology, 1(2), 317-355.
Gueimonde, M., Sakata, S., Kalliomäki, M., Isolauri, E., Benno, Y., & Salminen, S. (2006). Effect of maternal consumption of lactobacillus GG on transfer and establishment of fecal bifidobacterial microbiota in neonates. Journal of pediatric gastroenterology and nutrition, 42(2), 166-170.
Hardcastle, T. J., & Kelly, K. A. (2010). baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. BMC bioinformatics, 11(1), 422.
Heijtz, R., Wang, S., Anuar, F., Qian, Y., Björkholm, B., Samuelsson, A., Hibberd, M., Forssberg, H., and Pettersson, S. (2011). Normal gut microbiota modulates brain development and behavior. Proc. Natl. Acad. Sci. 108, pp. 3047–3052.
Hooper, L. V., Midtvedt, T., & Gordon, J. I. (2002). How host-microbial interactions shape the nutrient environment of the mammalian intestine. Annual review of nutrition, 22(1), 283-307.
Huson, D. H., Mitra, S., Ruscheweyh, H. J., Weber, N., & Schuster, S. C. (2011). Integrative analysis of environmental sequences using MEGAN4. Genome research, 21(9), 1552-1560.
Hyslop, E. J. (1980). Stomach contents analysis—a review of methods and their application. Journal of fish biology, 17(4), 411-429.
Ingerslev, H. C., von Gersdorff Jørgensen, L., Strube, M. L., Larsen, N., Dalsgaard, I., Boye, M., & Madsen, L. (2014). The development of the gut microbiota in rainbow trout (Oncorhynchus mykiss) is affected by first feeding and diet type. Aquaculture, 424, pp. 24-34.
Kang, D. D., Froula, J., Egan, R., & Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ, 3, e1165.
Kanther M. and Rawls JF. (2010). Host-microbe interactions in the developing zebrafish. Curr Opin Immunol 22, pp. 10–19.
Kikuchi, Y., Hayatsu, M., Hosokawa, T., Nagayama, A., Tago, K., & Fukatsu, T. (2012). Symbiont-mediated insecticide resistance. Proceedings of the National Academy of Sciences, 109(22), 8618-8622.
Koenig, J. E., Spor, A., Scalfone, N., Fricker, A. D., Stombaugh, J., Knight, R., ... & Ley, R. E. (2011). Succession of microbial consortia in the developing infant gut microbiome. Proceedings of the National Academy of Sciences, 108(Supplement 1), 4578-4585.
Kostic, A. D., Gevers, D., Siljander, H., Vatanen, T., Hyötyläinen, T., Hämäläinen, A. M., ... & Lähdesmäki, H. (2015). The dynamics of the human infant gut

microbiome in development and in progression toward type 1 diabetes. Cell host & microbe, 17(2), 260-273.
Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and environmental microbiology, 79(17), 5112-5120.
Larsen, N., Vogensen, F. K., van den Berg, F. W., Nielsen, D. S., Andreasen, A. S., Pedersen, B. K., ... & Jakobsen, M. (2010). Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. PloS one, 5(2), e9085.
Lawson, P. W., Logerwell, E. A., Mantua, N. J., Francis, R. C., & Agostini, V. N. (2004). Environmental factors influencing freshwater survival and smolt production in Pacific Northwest coho salmon (Oncorhynchus kisutch). Canadian Journal of Fisheries and Aquatic Sciences, 61(3), 360-373.
Lee WJ, Hase K. (2014). Gut microbiota-generated metabolites in animal health and disease. Nat Chem Biol., 10(6), pp. 416-24.
Llewellyn MS, Boutin S, Hoseinifar SH, Derome N. (2014). Teleost microbiomes: the state of the art in their characterization, manipulation and importance in aquaculture and fisheries. Front Microbiol. 5, 207.
Luczkovich, J. 1988. The role of prey detection in the selection of prey by the pinfish, Lagodon rhomboides (Linnaeus). Journal of Experimental Marine Biology and Ecology. 123: 15-30.
Luczkovich, Joseph J., Stephen F. Norton, and R. Grant Gilmore, Jr. 1995. The influence of oral anatomy on prey selection during ontogeny of two percoid fishes, Lagodon rhomboides and Centropomus undecimalis. Environmental Biology of Fishes, Vol. 44: 79-95.
Luczkovich, Joseph J. and E. J. Stellwag. 1993. Isolation of cellulolytic microbes from the intestinal tract of the pinfish, Lagodon rhomboides: size-related changes in diet and microbial abundance. Marine Biology 116: 389-379.
Magoč, T., & Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics, 27(21), 2957-2963.
Martin-Antonio, B., Manchado, M., Infante, C., Zerolo, R., Labella, A., Alonso, C., & Borrego, J. J. (2007). Intestinal microbiota variation in Senegalese sole (Solea senegalensis) under different feeding regimes. Aquaculture research, 38: pp. 1213-1222.
McCauley, D. J., Micheli, F., Young, H. S., Tittensor, D. P., Brumbaugh, D. R., Madin, E. M., ... & Worm, B. (2010). Acute effects of removing large fish from a near-pristine coral reef. *Marine biology*, 157, pp. 2739.
Metzker, M. L. (2010). Sequencing technologies—the next generation. Nature reviews genetics, 11(1), 31.
Miyake, S., Ngugi, D. K. and Stingl, U. (2015), Diet strongly influences the gut microbiota of surgeonfishes. Molecular Ecology, 24, pp. 656–672.
Mumby P. J., Dahlgren, C. P., Harborne, A. R., Kappel, C. V., Micheli, F., Brumbaugh, D. R., ... & Gill, A. B., (2006). Fishing, trophic cascades, and the process of grazing on coral reefs. *Science* 311, p. 98.
Myles, I. A., Fontecilla, N. M., Janelsins, B. M., Vithayathil, P. J., Segre, J. A., & Datta, S. K. (2013). Parental dietary fat intake alters offspring microbiome and

immunity. The Journal of Immunology, 191(6), 3200-3209.
Nacci, D., Coiro, L., Champlin, D., Jayaraman, S., McKinney, R., Gleason, T. R., ... & Cooper, K. R. (1999). Adaptations of wild populations of the estuarine fish Fundulus heteroclitus to persistent environmental contaminants. Marine Biology, 134(1), 9-17.
Nayak SK (2010). Role of gastrointestinal microbiota in fish. Aquaculture Research, 41, pp. 1553–1573.
Ott, S. J., Musfeldt, M., Wenderoth, D. F., Hampe, J., Brant, O., Fölsch, U. R., ... & Schreiber, S. (2004). Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. Gut, 53(5), 685-693.
Ottesen, E. A., Young, C. R., Eppley, J. M., Ryan, J. P., Chavez, F. P., Scholin, C. A., & DeLong, E. F. (2013). Pattern and synchrony of gene expression among sympatric marine microbial populations. Proceedings of the National Academy of Sciences, 110(6), E488-E497.
Ottesen, E. A., Young, C. R., Gifford, S. M., Eppley, J. M., Marin, R., Schuster, S. C., ... & DeLong, E. F. (2014). Multispecies diel transcriptional oscillations in open ocean heterotrophic bacterial assemblages. Science, 345(6193), 207-212.
Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H. Y., Cohoon, M., ... & Fonstein, M. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic acids research, 33(17), 5691-5702.
Pinckney, J. L., Örnólfsdóttir, E. B., & Lumsden, S. E. (2002). Estuarine phytoplankton group-specific responses to sublethal concentrations of the agricultural herbicide, atrazine. Marine Pollution Bulletin, 44(10), 1109-1116.
Poretsky, R. S., Hewson, I., Sun, S., Allen, A. E., Zehr, J. P., & Moran, M. A. (2009). Comparative day/night metatranscriptomic analysis of microbial communities in the North Pacific subtropical gyre. Environmental microbiology, 11(6), 1358-1375.
Porta, M., & Zumeta, E. (2002). Implementing the Stockholm treaty on persistent organic pollutants. Occupational and environmental medicine, 59(10), 651-652.
Rawls J., Samuel, B. S., & Gordon, J. I., (2004). Gnotobiotic zebrafish reveal evolutionarily conserved responses to the gut microbiota. Proc Natl Acad Sci, 101, pp. 4596–4601.
Roeselers G, Mittge EK, Stephens WZ et al. (2011). Evidence for a core gut microbiota in the zebrafish. ISME Journal, 5, 1595–1608.
Round, J. and Mazmanian S. (2009). The gut microbiota shapes intestinal immune responses during health and disease. Nat Rev Immunol, 9, pp. 313–323.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proceedings of the national academy of sciences, *74*(12), 5463-5467.
Schmidt, V.T., Smith, K.F., Melvin, D.W., & Amaral-Zettler, L.A. (2015). Community assembly of a euryhaline fish microbiome during salinity acclimation. Molecular Ecology, 24, pp 2537-2550.
Senderovich, Y., & Halpern, M. (2013). The protective role of endogenous bacterial communities in chironomid egg masses and larvae. The ISME journal, 7(11),

2147-2158.
Sharma, R., Young, C., & Neu, J. (2010). Molecular modulation of intestinal epithelial barrier: contribution of microbiota. BioMed Research International, 2010.
Sonnenburg, J. L., Xu, J., Leip, D. D., Chen, C. H., Westover, B. P., Weatherford, J., ... & Gordon, J. I. (2005). Glycan foraging in vivo by an intestine-adapted bacterial symbiont. Science, 307(5717), 1955-1959.
Spanggaard, B., I. Huber, J. Nielsen, T. Nielsen, K. F. Appel, and L. Gram. (2000). The microflora of rainbow trout intestine: a comparison of traditional and molecular identification. Aquaculture, 182, pp. 1-15.
Spor, A., Koren, O., & Ley, R. (2011). Unravelling the effects of the environment and host genotype on the gut microbiome. Nature Reviews Microbiology, 9(4), 279-290.
Stellwag, E.J. T. D. Smith, and J. J. Luczkovich. 1995. Characterization and ecology of carboxymethycellulase-producing anaerobic bacterial communities associated with the intestinal tract of the pinfish, Lagodon rhomboides. Applied and Environmental Microbiology, February 1995, 813-816.
Stephens, W. Z., Burns, A. R., Stagaman, K., Wong, S., Rawls, J. F., Guillemin, K., & Bohannan, B. J. (2015). The composition of the zebrafish intestinal microbial community varies across development. The ISME journal.
Stewart JA, Chadwick VS, Murray A (2005) Investigations into the influence of host genetics on the predominant eubacteria in the faecal microflora of children.
Sullam, K.E., Essinger, S. D., Lozupone, C. A., O'CONNOR, M. P., Rosen, G. L., Knight, R. O. B., ... & Russell, J. A. (2012). Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. Molecular Ecology, 21, pp. 3363-3378.
Sullam, K. E., Rubin, B. E., Dalton, C. M., Kilham, S. S., Flecker, A. S., & Russell, J. A. (2015). Divergence across diet, time and populations rules out parallel evolution in the gut microbiomes of Trinidadian guppies. The ISME journal.
Theis KR, Venkataraman, A., Dycus, J. A., Koonter, K. D., Schmitt-Matzen, E. N., Wagner, A. P., ... & Schmidt, T. M. (2013). Symbiotic bacteria appear to mediate hyena social odors. Proc Natl Acad Sci., 110, 49, pp. 19832–19837.
Turnbaugh, P.J., Ley, R. E., Mahowald, M. A., Magrini, V., Mardis, E. R., & Gordon, J. I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest, Nature, 444, pp. 1027-1031.
Turnbaugh, P. J., Bäckhed, F., Fulton, L., & Gordon, J. I. (2008). Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. Cell host & microbe, 3(4), 213-223.
Turnbaugh, P. J., Hamady, M., Yatsunenko, T., Cantarel, B. L., Duncan, A., Ley, R. E., ... & Egholm, M. (2009). A core gut microbiome in obese and lean twins. nature, 457(7228), 480-484.
Vaishampayan, P. A., Kuehl, J. V., Froula, J. L., Morgan, J. L., Ochman, H., & Francino, M. P. (2010). Comparative metagenomics and population dynamics of the gut microbiota in mother and infant. Genome biology and evolution, 2, 53-66.
Voreades, N., Kozil, A., & Weir, T. L. (2015). Diet and the development of the human intestinal microbiome. Recent Discoveries in Evolutionary and Genomic Microbiology, 40.

Wada-Katsumata, A., Zurek, L., Nalyanya, G., Roelofs, W. L., Zhang, A., & Schal, C. (2015). Gut bacteria mediate aggregation in the German cockroach. Proceedings of the National Academy of Sciences, 112(51), 15678-15683.

Whitman WB, Coleman DC, Wiebe WJ (1998). Prokaryotes: the unseen majority. Proc Natl Acad Sci 95, pp. 6578–6583.

Xing, M., Hou, Z., Yuan, J., Liu, Y., Qu, Y., & Liu, B. (2013). Taxonomic and functional metagenomic profiling of gastrointestinal tract microbiome of the farmed adult turbot (Scophthalmus maximus). FEMS Microbiology Ecology, 86(3), pp. 432-443.

Ye L, Amberg J, Chapman D, Gaikowski M, Liu WT. (2014). Fish gut microbiota analysis differentiates physiology and behavior of invasive Asian carp and indigenous American fish. ISME J., 8, pp. 541–551.

Zheng X, Zhao, A., Xie, G., Chi, Y., Zhao, L., Li, H., ... & Jia, W. (2013). Melamine-induced renal toxicity is mediated by the gut microbiota. Sci Transl Med., 5, 172, 172ra22.

Zoetendal, E. G., Akkermans, A. D., Akkermans-van Vliet, W. M., de Visser, J. A. G., & de Vos, W. M. (2001). The host genotype affects the bacterial community in the human gastronintestinal tract. Microbial ecology in health and disease, 13(3), 129-134.

CHAPTER 2.

MICROBIAL EUKARYOTE DIVERSITY IN THE MARINE OXYGEN MINIMUM ZONE OFF NORTHERN CHILE

Disclaimer: This chapter was published with the same title in the journal Frontiers in Microbiology on 28 October, 2014.

Citation: Parris, D. J., Ganesh, S., Edgcomb, V. P., DeLong, E. F., & Stewart, F. J. (2014). Microbial eukaryote diversity in the marine oxygen minimum zone off northern Chile. *Frontiers in microbiology*, *5*, 543

2.1 Abstract

Molecular surveys are revealing diverse eukaryotic assemblages in oxygen-limited ocean waters. These communities may play pivotal ecological roles through autotrophy, feeding, and a wide range of symbiotic associations with prokaryotes. We used 18S rRNA gene sequencing to provide the first snapshot of pelagic microeukaryotic community structure in two cellular size fractions (0.2-1.6 µm, >1.6 µm) from seven depths through the anoxic oxygen minimum zone (OMZ) off northern Chile. Sequencing of >154,000 amplicons revealed contrasting patterns of phylogenetic diversity across size fractions and depths. Protist and total eukaryote diversity in the >1.6 µm fraction peaked at the chlorophyll maximum in the upper photic zone before declining by ~50% in the OMZ. In contrast, diversity in the 0.2-1.6 µm fraction, though also elevated in the upper photic zone, increased four-fold from the lower oxycline to a maximum at the anoxic OMZ core. Dinoflagellates of the Dinophyceae and endosymbiotic Syndiniales clades dominated the protist assemblage at all depths (~40-70% of sequences). Other protist groups varied with depth, with the anoxic zone community of the larger size fraction enriched in euglenozoan flagellates and acantharean radiolarians (up to 18% and 40% of all sequences, respectively). The OMZ 0.2-1.6 µm fraction was dominated (11-99%) by Syndiniales, which exhibited depth-specific variation in composition and total richness despite uniform oxygen conditions. Metazoan sequences, though confined primarily to the 1.6 µm fraction above the OMZ, were also detected within the anoxic zone where groups such as copepods increased in abundance relative to the oxycline and upper OMZ. These data, compared to those from other low-oxygen sites, reveals variation in OMZ

microeukaryote composition, helping to identify clades with potential adaptations to oxygen-depletion.

2.2 Introduction

Marine oxygen minimum zones (OMZs) support complex microbial communities adapted for life under low oxygen conditions. In the anoxic OMZ of the Eastern Tropical South Pacific (ETSP) off Chile and Peru, aerobic respiration of sinking organic matter combines with a decline in water column mixing to deplete dissolved oxygen to below detection (<15 nM; Thamdrup et al. 2012). Oxygen depletion at the ETSP OMZ core (~100-400 m) selects for communities dominated by bacteria and archaea (prokaryotes) capable of anaerobic or microaerophilic metabolism (Stevens and Ulloa, 2008, Ulloa et al. 2012). These microorganisms make important contributions to marine nitrogen, sulfur, and carbon cycling (Wright et al. 2012), for example as mediators of nitrogen loss through denitrification and anaerobic ammonium oxidation (Thamdrup et al. 2006, Lam et al. 2009). Although most large multicellular eukaryotes are absent from anoxic marine waters, severely oxygen-depleted layers in other OMZ regions contain a diverse community of microbial eukaryotes (microeukaryotes) including protists and fungi, as well as zooplankton with potential adaptations to oxygen limitation (Orsi and Edgcomb 2013, Orsi et al. 2012; Teuber et al. 2013). As suggested for a coastal OMZ (Orsi et al 2012), microeukaryotes in the ETSP OMZ may play important ecological roles, for example through predation or symbiotic interactions with prokaryotes. However, while bacteria and archaea in the ETSP OMZ are becoming better understood, and numerous studies have examined benthic meiofaunal diversity beneath OMZs (Neira et al. 2001,

Levin 2003), little is known about pelagic eukaryote diversity in this unique oxygen-depleted environment.

Marine microeukaryotes and zooplankton can significantly affect the structure and function of prokaryote communities and are now recognized as pivotal members of aquatic food webs in numerical models of nutrient cycling and in paradigms of surface and deep-ocean ecology (Arístegui et al. 2009). Protists impact major nutrient cycles directly and indirectly, through grazing on prokaryotic prey and consequent regeneration of nutrients, and through direct modification of organic matter (sinking particulate and dissolved) (Taylor, 1982; 1986; Tang et al. 2010; Sherr and Sherr 2002). Zooplankton surfaces, guts, and fecal pellets are habitats for bacterial growth (Tang et al. 2010), supporting local cell densities orders of magnitude higher than in the surrounding water (Grossart 2003). Studies in oxygen deficient waters of the North Pacific and the anoxic Cariaco Basin have revealed unique protist communities with diversity and abundances varying along gradients of oxygen, sulfide, and methane concentration (Taylor et al. 2001; Li et al. 2008; Lin et al. 2008; Edgcomb et al. 2011; Orsi et al. 2011, 2012). Anaerobic and microaerophilic species of ciliates and euglenid flagellates seem to be enriched in anoxic water layers (Lynn 2008; Edgcomb et al. 2011, Orsi et al 2011, Orsi et al. 2012). Recent studies of marine sediments suggest that some euglenids engage in symbiosis with nitrate reducing, sulfur oxidizing, epibiotic bacteria (Bernhard et al. 2000; Yubuki et al. 2009; Edgcomb et al. 2010). Cryptic symbioses between ciliates and methanogenic archaea have been suggested to possibly contribute significantly to methanogenesis in OMZs (Orsi et al. 2012). Furthermore, OMZs are known to harbor larger zooplankton such as copepods with specific ecophysiological and behavioral

adaptations for persistence under low oxygen (e.g., reduced overall metabolic rates, enhanced activity of anaerobic pathways, diel migrations) (Childress and Seibel 1998, Teuber et al. 2013). However, interactions of OMZ zooplankton with other community members are not well understood.

In some ocean regions, picoeukaryotes smaller than 3 µm may be among the most abundant cells in the water column (Moon-van der Staay, 2001, Massana 2011), playing roles in primary production, food web dynamics, and mineral cycling (Fogg, G.E. 1995, Diez et al. 2001, Moon-van der Staay, 2001, Worden et al. 2004). However, many of these smaller microeukaryotes are new to science, and we are just beginning to understand their ecology (Moon-van der Staay, 2001, Worden et al. 2004). Because of their small size, picoeukaryotes may be overlooked in gene surveys focusing only on larger particulate size fractions of filtered water samples (>3.0 µm). Indeed, no studies of OMZs have presented a size-fractionated comparison of eukaryotic community structure. Characterizing microeukaryote diversity and distribution among different OMZ regions and size fractions is a necessary step toward quantifying the ecological importance of these organisms for bulk chemical cycling and eukaryote-prokaryote interactions in OMZs and for clarifying the effects of severe oxygen depletion on planktonic community structure.

Here, we use 18S rRNA gene amplicon sequencing to provide a snapshot of eukaryote community structure in two size fractions (0.2-1.6 µm, >1.6 µm) along a depth gradient through the ETSP OMZ off northern Chile. To our knowledge, this is the first community-level genetic survey of microbial eukaryotes in this anoxic water column.

We highlight differences between the two size fractions, across the OMZ oxygen gradient, and make comparisons to 18S gene surveys from other marine environments.

2.3 Materials and Methods

2.3.1 Sample collection and DNA extraction

Microbial community samples were collected from the ETSP OMZ as part of the Center for Microbial Ecology: Research and Education (C-MORE) BiG RAPA (Biogeochemical Gradients: Role in Arranging Planktonic Assemblages) cruise aboard the *R/V Melville* (18 November - 14 December, 2010). Seawater was sampled from 7 depths spanning the oxic photic zone and oxycline (5 m and 32 m samples), the suboxic (O_2 < 10 µm) upper OMZ (70 m), the anoxic OMZ core (110, 200, 320 m), and the oxic zone beneath the OMZ (1000 m) from a single site (20° 05.0 S, 70° 48.0 W) off the coast of Iquique, Chile on November 19th (5m), 20th (32m), 21st (70, 110, 200, 320 m) and 23rd (1000 m). Water was sampled from Niskin bottles on a rosette containing a CTD profiler (Sea-Bird SBE 911plus) with a dissolved oxygen sensor and fluorometer. A peristaltic pump was used to filter water (10 L) through an in-line prefilter (GF/A, 1.6 µm pore-size, 47mm dia., Whatman) to retain larger eukaryotes and a Sterivex™filter (0.22 µm pore-size, Millipore) to retain picoeukaryotes. Prefilters and Sterivex™ filters were saturated with lysis buffer (~1.8 ml; 50 mM Tris-HCl, 40 mM EDTA, and 0.73 M Sucrose) and frozen until analysis.

Genomic DNA was extracted as described in Ganesh et al. (2014). Cells were lysed by adding lysozyme (2 mg in 40 µl of Lysis buffer per filter) directly to a microcentrifuge tube containing a pre-filter or to the Sterivex™ cartridge, sealing the caps/ends, and incubating for 45 min at 37°C. Proteinase K (1 mg in 100 µl lysis buffer,

with 100 µl 20% SDS) was added and the tubes further incubated for 2 hours at 55ºC. Lysate was removed and DNA was extracted once with Phenol:Chloroform:Isoamyl Alcohol (25:24:1) and once with Chloroform:Isoamyl Alcohol (24:1). The purified aqueous phase was concentrated using Amicon Ultra-4 w/100 kDa MWCO centrifugal filters. Purified DNA aliquots from each depth and size fraction were used for PCR.

2.3.2 18S rRNA gene PCR and amplicon sequencing

A 450-bp fragment spanning the variable V4 region of the 18S rRNA gene was PCR amplified using the universal forward primer 3NDF and the universal reverse primer V4_euk_R1 (Brate et al. 2010). Primers were appended with Roche 454 sequencing adaptors and sample-specific barcodes as described in the protocol established for the Human Microbiome Project by the Broad Institute (Jumpstart Consortium Human Microbiome Project Data Generation Working Group, 2012). Reactions (25 µl) were performed with 50-150 ng template DNA in Platinum PCR Mastermix (Invitrogen) and run on an Eppendorf Mastercyler thermocycler. The PCR program for amplification was: 94 °C for 2 min, followed by 34 cycles of 30 s at 94 °C, 30 s at 59 °C, 60 s at 72 °C with a final extension at 72 °C for 7 min. Amplicons were purified using the QIAQuick PCR Clean-Up Kit, quantified on a Qubit® 2.0 Fluorometer, and pooled at equimolar concentrations. Pooled, barcoded amplicons were sequenced unidirectionally using the Roche Genome Sequencer FLX Instrument with Titanium chemistry. All sequence data generated in this study will be made publicly available in the NCBI Sequence Read Archive upon publication. Data are available for reviewers upon request.

2.3.3 Data Analysis

Amplicons were analyzed according to established protocols using the QIIME package (Caporaso et al. 2010). Barcoded 18S datasets were de-multiplexed and filtered to remove low quality sequences using default parameters (minimum quality score = 25, minimum sequence length = 200, no ambiguous bases allowed). De-multiplexed sequences were clustered into Operational Taxonomic Units (OTUs) at 97% sequence similarity, with taxonomy assigned to representative OTUs from each cluster by queries against the Protist Ribosomal Reference database, PR2 (as of March 2014, Guillou et al. 2010). OTU counts were rarefied and alpha diversity was quantified at a uniform sequencing depth based on the lowest sequence count from the protist only subset at 70m (N = 1308 sequences) using the phylogenetic diversity (PD) metric (Faith 1992). To evaluate differences in community composition between samples, sequences were aligned using the PyNAST aligner in QIIME and beta diversity was calculated using the weighted Unifrac metric, which compares samples based on the phylogenetic relatedness (branch lengths) of OTUs in a community, taking into account relative OTU abundance (Lozupone and Knight, 2005). Sample relatedness based on Unifrac was visualized using a two-dimensional Principal Coordinate Analysis.

2.4 Results

2.4.1 Oxygen

As reported previously (Ganesh et al. 2014), the sample site was characterized by steep vertical gradients in dissolved oxygen (Figure 2.1). Oxygen concentrations measured with the SBE sensor ranged from ~250 µM at the surface to below 5 µM through the OMZ core (~100-400 m), before increasing below 400 m to ~60 µM at 1000 m. Prior measurements using high sensitivity switchable trace oxygen sensors with

resolution in the nanomolar range indicate that the ETSP OMZ core is anoxic, with oxygen concentrations below 30 nM (Revsbech et al. 2009, Thamdrup et al. 2012). The base of the photic zone (1% surface PAR) occurred at ~40 m, with chlorophyll concentration peaking at 32 m (1.97 µg l^{-1}).

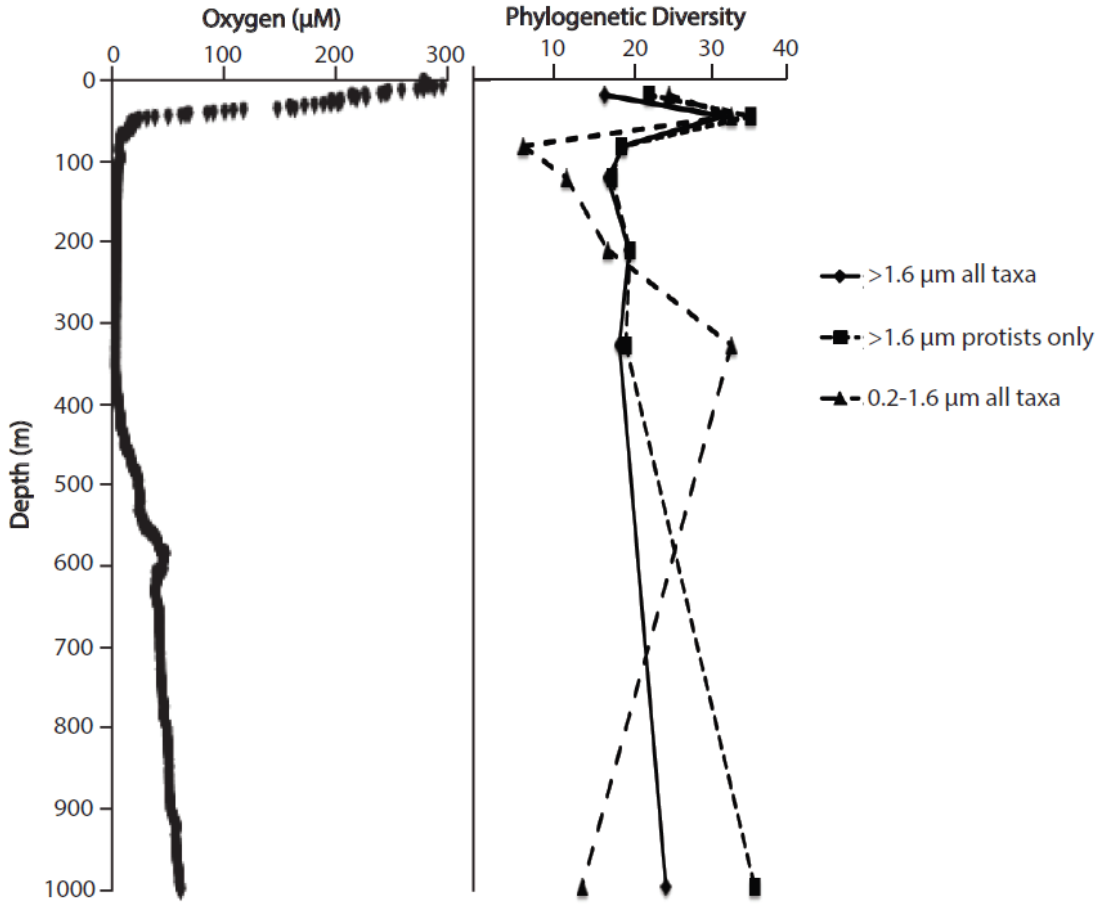

Figure 2.1. Dissolved oxygen concentration and phylogenetic diversity as a function of water column depth. Diversity data points are mean values based on rarefaction of OTU counts at a standardized sequence count (n = 1308 per sample).

Table 2.1. 18S rRNA gene amplicon sequencing statistics

Sample	Count	OTU[1]
TOTAL sequences	196,027[2]	
Mean length	410 bp	
TOTAL unique OTUs		2331
5m 0.2-1.6 µm	4,679	463
5m >1.6 µm	26,127	639
32m 0.2-1.6 µm	6,760	620
32m >1.6 µm	12,943	946
70m 0.2-1.6 µm	2,036	47
70m >1.6 µm	15,564	462
110m 0.2-1.6 µm	8,097	205
110m >1.6 µm	14,868	440
200m 0.2-1.6 µm	5,633	270
200m >1.6 µm	13,388	504
320m 0.2-1.6 µm	7,148	692
320m >1.6 µm	9,562	382
1000m 0.2-1.6 µm	10,622	310
1000m >1.6 µm	16,689	660

[1] observed number of operational taxonomic units (>97% nucleotide similarity)
[2]Total sequences before quality filtering

2.4.2 Phylogenetic diversity

Following removal of low quality reads, 44,975 and 109,141 sequences representing both protist and metazoan taxa were analyzed to assess diversity in the 0.2-1.6 µm (picoeukaryote) and >1.6 µm size fractions, respectively (median per sample: 6,760 and 15,564 for small and large size fractions; Table 2.1, Figure 2.1). In the >1.6 µm fraction, phylogenetic diversity (PD), which measures the total branch length connecting all OTUs in the 18S rRNA gene phylogeny, peaked at 32 m within the primary chlorophyll maximum, but was comparable between surface (5 m) and anoxic (110-320 m) depths (Figure 2.1). In contrast, PD of the picoeukaryote fraction was bimodal with depth, peaking first within the chlorophyll maximum layer, declining to a

minimum along the oxycline, and then progressively increasing four-fold within the OMZ to peak again at 300 m, where diversity was equivalent to that observed at 32 m and almost two-fold higher than that of the >1.6 μm fraction at the same depth (Figure 2.1).

2.4.3 Community composition

Eukaryote community composition differed substantially along the OMZ oxygen gradient and between size fractions (Figure 2.2). Based on weighted Unifrac distances, sequences of the >1.6 μm fraction were less similar across depths compared to the 0.2-1.6 μm fraction (Figure 2.2E), reflecting a broader range of major taxonomic groups in the larger fraction. Indeed, excluding the 5 m surface sample, samples of the 0.2-1.6 μm fraction were dominated by a single Order (Syndiniales). The >1.6 μm fraction from the oxic sub-OMZ (1000 m) and both size fractions from the surface (5 m) sample were relative outliers, clustering apart from those of the oxycline and OMZ depths (Figure 2.2E).

The separation of the surface samples was due partly to their enrichment in metazoan and haptophyte sequences (45-72% of total, compared to 3-13% for sub-surface depths; Figure 2.2C). The diversity of the metazoan pool peaked at 5 m (Figure 2.2D), with groups such as the Appendicularia (Urochordata) and arthropods over-represented relative to OMZ depths (Figure 2.2D). In contrast, cnidarians and ctenophores comprised the bulk of the metazoan sequences in the larger size fraction within and below the OMZ. Sequences matching copepods (Maxillopoda) declined in relative abundance from the primary chlorophyll layer (32 m) to the upper OMZ, but increased again to a local maximum at the OMZ core (200, 320 m; Figure 2.2D).

Protists were abundant (45-96%) at all depths but comprised the vast majority (>85% of total) of the eukaryote community within the oxycline and OMZ (Figure 2.2B). Several groups followed clear depth-specific trends. Euglenozoan flagellates of the class Diplonemea represented a major portion of the OMZ community, comprising 18% of total sequences from the larger size fraction at 200 m, but were almost absent (<2%) from oxycline and surface communities. Ciliate sequences related to the Litostomatea were only observed in the 110 m and 200 m OMZ depths, although in low abundance (0.5-1.5% of all sequences). Photosynthetic algae of the Cryptophyta, Haptophyta, Picobiliphyta, Choanoflagellida, and Stramenopiles (diatoms) were confined primarily to the surface, representing ~20% of the sequence pool from the larger size fraction at 5 m (Figure 2.2B). In contrast, radiolarians in this fraction became progressively more dominant with depth, with unidentified members of the Acantharea peaking at the upper and lower OMZ boundaries (110m and 320m), accounting for 42% and 41% of sequences, respectively, and Polycistenea radiolarians peaking at over 1/3 of protist sequences at 1000 m beneath the OMZ (Figure 2.2B).

The most consistently abundant protist OTUs were similar to uncultured Dinophyceae and the predominantly parasitic dinoflagellate taxon Syndiniales. These groups were detected at all depths, representing 42-68% of all protist sequences in the larger size fraction. Aside from a decline in Dinophyceae representation at 1000 m (to ~7%), sequence abundance within Dinophyceae and Syndiniales in the larger size fraction did not differ notably between oxic and OMZ depths (Figure 2.2B). In contrast, the picoeukaryotic fraction (excluding that of the surface sample) was dominated by Syndiniales dinoflagellates (53-99% of total), which showed substantial depth-specific

variation in clade diversity across depths (Figure 2.2F). Syndiniales diversity peaked at the surface and again within the OMZ, with Group II clades 5, 21, and 30 enriched at the OMZ depths (Figure 2.2E) and Group II clade 7 being the most abundant group at all other depths.

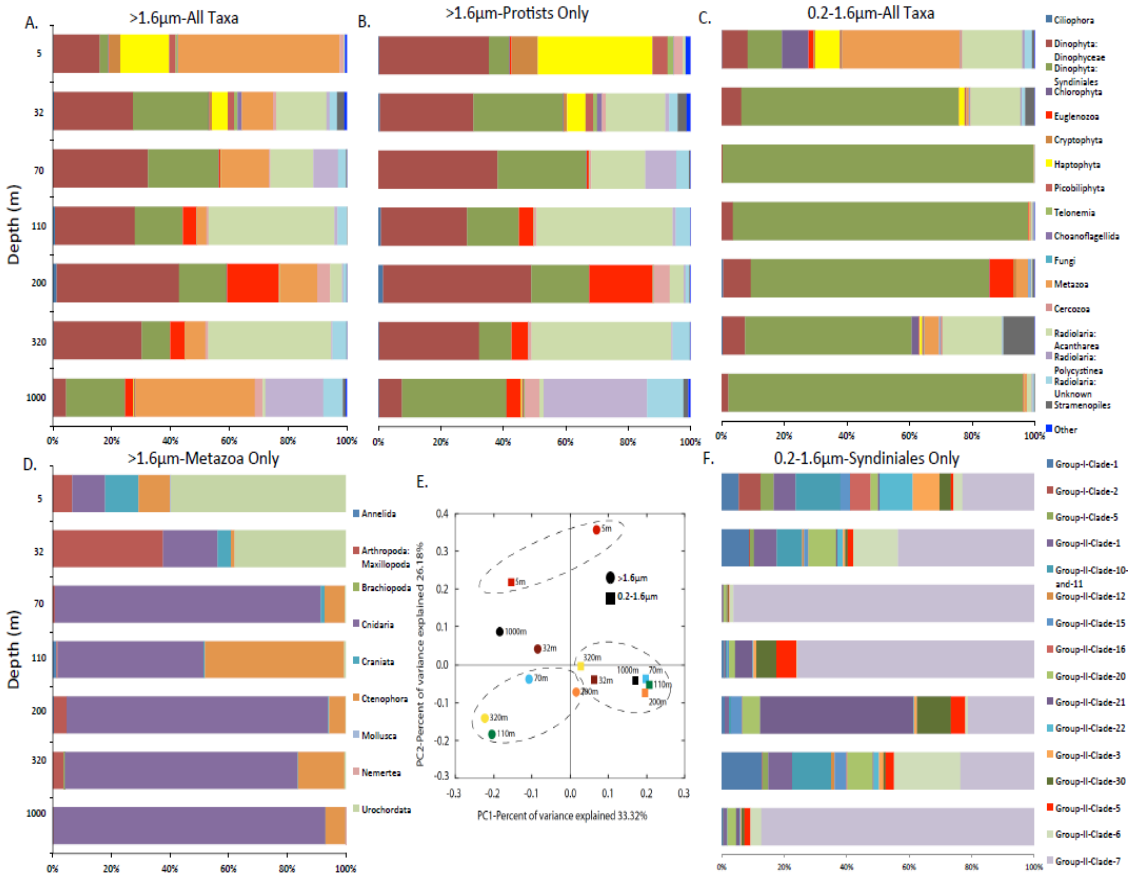

Figure 2.2. Community taxonomic composition in the (A) >1.6 µm fraction based on all 18S rRNA gene sequences, (B) >1.6 µm fraction with metazoan sequences excluded, (C) 0.2-1.6 µm fraction based on all sequences, (D) >1.6 µm fraction based only on sequences matching metazoa, (E) shows a principle component analysis of community taxonomic relatedness based on all sequences from both size fractions, as quantified by the weighted Unifrac metric. Depths within the OMZ are circled. (F) 0.2-1.6 µm size fraction based only on sequences matching Syndiniales dinoflagellates.

Table 2.2: Relative abundance of dominant taxa across diverse low oxygen sites

Site	Oxygen conditions[a]	Taxanomic group[b]	Relative %	Reference	Method[c]	# of sequences
North Pacific Ocean	Oxic at 5m	Ciliophora	20%	Countway et al. 2010	CS	524
Oxic site, January sampling		Unknown Alveolate	20%			
		Dinophysceae	16%			
	Oxic at 150m	Polycistenea	70%			
		Marine Alveolate Group I	12%			
		Acantharea	3%			
	Oxic at 500m	Polycistenea	55%			
		Marine Alveolate Group II	23%			
		Marine Alveolate Group I	7%			
Hamelin Pool, Shark Bay microbialites, Eastern Mediterranean	Anoxic at 1-2cm	Litostomatea (Ciliophora)	40%	Edgcomb and Bernhard, 2013	A, CS	>96
Permanent anoxic basin		Labyrinthulida (Stramenopile)	10-20%			
		Foraminifera (Rhizaria)	10-20%			
	Oxic at 0-1cm	Dinophysceae (Protodinium, Hetercapsaceae)	10-20%			
		Stramenopiles (Labryinthulida)	10-20%			
Cariaco Basin, Caribbean Sea	Oxic	Radiolarians (RAD-3, RAD-19)	90%	Orsi et al. 2011	CS	6489
Permanent anoxic basin		MAST Stramenopiles	5%			
		Novel Alveolates	5%			
	Anoxic	Radiolarians (RAD-3, RAD-19)	72%			
		MAST Stramenopiles	11%			
		Novel Alveolates	14%			
		Euglenozoa	1.50%			
	Anoxic/Sulfidic	Radiolarians (RAD-3, RAD-19)	60%			
		MAST Stramenopiles	17%			

Table 2.2 continued

		Novel Alveolates	14%			
		Euglenozoa	9%			
Indian Ocean						
Oxic site	Oxic at surface	Marine alveolates I	20-35%	Not et al. 2008	CS	541
		Dinophysceae	10-25%			
		MAST Stramenopiles	10-20%			
	Oxic at chl max	Marine alveolates I	10-30%			
		Radiolarians	15-40%			
		MAST Stramenopiles	5-15%			
Saanich Inlet	Oxic	Stramenopiles	40-50%	Orsi et al. 2012	A	4987
Seasonally anoxic fjord		Dinophysceae	20-30%			
	Suboxic	Stramenopiles	30-40%			
		Dinophysceae	30-40%			
	Anoxic	Stramenopiles	15-25%			
		Ciliophora	15-20%			
		Dinophysceae	15-20%			
Saanich Inlet	Anoxic (200m)	Dinophysceae	21%	Unpublished data	A	2000
Seasonally anoxic fjord		Dinophyta (Syndininales)	17%	Edgcomb et al.		
		Stramenopiles	15%			
Gulf of Mexico	Oxic	Coscinodiscophyceae	39%	Rocke et al. 2012	CS	175
Seasonally anoxic bottom waters		Dinophysceae	40-50%			
		Prasinophytes	10%			
	Anoxic	Dinophysceae	80%			
		Polycystenea	9%			
		Novel Alveolates	6%			
North Pacific	Oxic	Ciliophora	27%	Schnetzer et al. 2011	CS	856

Table 2.2 continued

Coastal upwelling near California		Stramenopiles	20%			
		Dinophysceae	14%			
	Suboxic	Alveolate Group II	35%			
		Acantharea	19%			
		Polycystenea	16%			
	Anoxic	Alveolate Group II	36%			
		Ciliophora	16%			
		Acantharea	13%			
Gotland Deep, Baltic Sea	Suboxic	Ciliophora	43%	Stock et al. 2009	CS	600
Seasonally anoxic basin		Fungi	30%			
		Choanoflagellida	11%			
	Anoxic/Sulfidic	Jakobida	71%			
		Fungi	6%			
		Ciliophora	5%			
Thetis, Mediterranean Sea	Anoxic	Fungi	37%	Stock et al. 2012	CS	192
Hypersaline, deep, anoxic basin		Ciliophora	20%			
		Stramenopiles	20%			
Sippewissett Salt Marsh, Cape Cod, Mass	Suboxic	Unknown Stramenopiles	26%	Stoeck and Epstein 2003	CS	42
Suboxic water column above sediments		Unkown Alveolates (mainly Ciliophora)	26%			
Black Sea	Suboxic	Stramenopiles	42%	Wylezich and Jurgens, 2011	CS	258
Permanently anoxic portions of water column		Ciliophora	36%			
		Dinoflagellates	19%			
	Anoxic/Sulfidic	Stramenoplies	42%			
		Ciliophora	33%			
		Euglenozoa	8%			

Table 2.2 continued

Mariajer Fjord, Denmark	Anoxic	Alveolates (mainly Ciliophora and Dinoflagellates)	41%	Zuendorf et al. 2006	CS	307
Permanently anoxic fjord		Stramenopiles	28%			
Eastern Tropical North Pacific OMZ	Oxic at 5m	Metazoa	55%	This Study	A	196,027
Permanently anoxic portions of water column		Dinophysceae	15%			
		Prymnesiophysceae	5%			
	Oxic at 32m	Dinophysceae	27%			
		Dinophyta (Syndininales)	26%			
		Acantharea	17%			
	Oxic at 70m	Dinophysceae	32%			
		Dinophyta (Syndininales)	24%			
		Metazoa	16%			
	Anoxic at 110m	Acantharea	43%			
		Dinophysceae	27%			
		Dinophyta (Syndininales)	16%			
	Anoxic at 200m	Dinophysceae	40%			
		Euglenozoa	18%			
		Dinophyta (Syndiniales)	16%			
	Anoxic at 320m	Acantharea	40%			
		Dinophysceae	25%			
		Dinophyta (Syndiniales)	8%			
	Oxic at 1000m	Metazoa	40%			
		Polycistenea	20%			
		Dinophyta (Syndininales)	20%			

[a] Lightly shaded areas represent low oxygen samples, Oxic= O2 above 15µM, Suboxic= O2 below 15µM, Anoxic or Anoxic/Sulfidic= undetectable O2

[b] Dominant taxa, darkly shaded taxa are those that are shared among multiple low oxygen sites

[c] Sequencing Method, CS= cloning and Sanger sequencing, A=Amplicon sequencing

2.5 Discussion

As observed in other low oxygen marine waters (Table 2.2), the ETSP OMZ harbors a diverse community of microeukaryotes with complex patterns of taxonomic structuring across oxygen gradients and between cellular size fractions. Phylogenetic diversity of eukaryotes in the >1.6 um size fraction, which presumably captures the majority of eukaryotic biomass, peaked in the oxic, photic zone in the layer of maximum chlorophyll fluorescence. This pattern is consistent with the hypothesis that resource availability supports a greater number of niches for microeukaryotes. Enhanced diversity in the photic zone may also be driven by a greater overall abundance in near-surface waters of larger organisms (metazoa) acting as hosts for symbiotic or parasitic protists. Interestingly, diversity within the larger size fraction from anoxic OMZ depths was equivalent to that at the surface (5 m) and in the oxic bottom water. The overall diversity trend in the larger size fraction did not vary substantially when only protists were analyzed (Figure 2.1), indicating that metazoan sequences had negligible effects on diversity calculations.

The relationship between picoeukaryotes and oxygen in OMZs is not well understood. In contrast to the larger size fraction, the 0.2-1.6 μm fraction in the ETSP was most diverse at the OMZ core. This pattern differs from that observed in seasonally anoxic waters where microeukaryote diversity (0.22-2.7 μm) and species richness declines as oxygen decreases below 20 mM (Orsi et al. 2012). However, the presence of sequences from metazoans and larger protists (e.g., radiolarians) in the 0.2-1.6 μm fraction suggests that a portion of the DNA in this size range likely originated from lysed cells and that diversity patterns in this fraction should be interpreted cautiously.

Nonetheless, depth-specific diversity trends in the picoeukaryote fraction do not directly parallel those of the larger size fraction. A similar decoupling was observed in a recent analysis of size-fractionated prokaryotic communities at this ETSP site where diversity decreased from the surface to anoxic waters in the small size fraction while increasing in the larger size fraction (Ganesh et al. 2014). Though limited to a single depth profile, our data indicate that relationships between environmental variables, notably oxygen concentration, and community diversity metrics vary depending on the subset of the community being examined, indicating that estimates of bulk eukaryote diversity and inferences of diversity-environment linkages in this system may require examination of multiple microbial size classes.

The composition of the OMZ eukaryote community in the larger size fraction was presumably strongly influenced by the community in the overlying oxic, photic zone. The majority of non-metazoan sequences in this fraction were found at the surface (5 m) and at the depth of maximum chlorophyll fluorescence (32 m), and included diverse algae, particularly photosynthetic haptophytes, cryptophytes, unaffiliated Dinophyceae, and Stramenopiles (diatoms), as well as heterotrophic radiolarians belonging to the Acantharea and putative endosymbiotic dinoflagellates of the Syndiniales. Although haptophyte and cryptophyte sequence abundance became negligible upon the transition out of the photic zone and into the OMZ, radiolarians, Dinophyceae and Syndiniales dinoflagellates remained dominant community members at all depths. The consistent signal from dinoflagellates may, to an unknown extent, reflect their presence as sinking dinospores resulting from near-surface infections, or alternatively, their presence in sinking, inactive or dead metazoan hosts. Similarly, Stramenopile sequences mostly

related to photosynthetic diatoms were detectable at all depths and in both size fractions, though at low abundance (<2%). Stramenopiles, diatoms in particular, are known to form dense blooms in upwelling zones, sink rapidly when no longer growing, and are major constituents of OMZ eukaryotic communities at the surface and most likely as sinking particulate organic matter at anoxic depths (Table 2.2, Smetacek, 1985, Montero et al. 2007). The low relative abundance of stramenopiles in this dataset may indicate that we sampled between blooms of these taxa or during a bloom of other algal species. We cannot rule out the possibility that some of the diatom sequences recovered from anoxic depths represent active cells. It has been shown, for example, that some diatoms may respire nitrate to persist during periods of dark anoxia (Kamp et al. 2011), raising the possibility that even transient (sinking) members of the OMZ microeukaryote assemblage may contribute to bulk elemental cycling in anoxic waters.

Although broad similarities in microeukaryote composition were observed across depths, several groups were enriched within the anoxic zone. Euglenid flagellates matching most closely to *Diplonema* and ciliates of the *Litostomatea* increased in relative abundance in the OMZ, particularly in the larger size fraction. Diverse euglenids and ciliates are common inhabitants of anaerobic and microaerophilic marine waters (Bernard and Fenchel 1996, Hayward et al. 2003, Orsi, et al. 2011, Table 2.2). Clone-based molecular studies of OMZs have identified a new class of euglenids (Symbiontida) that are closely related to diplonemids, common in low oxygen waters, and appear to host sulfide-oxidizing, nitrate reducing epibiotic bacteria (Edgcomb et al. 2010; Orsi et al. 2011, Orsi et al. 2012). Many different ciliates, including those in the class *Litostomatea*, have been recovered from molecular and visual studies of low oxygen sites (Lynn, 2008,

Stock et al. 2009, Edgcomb and Bernhard, 2013; Edgcomb and Pachiadaki 2014). Taxa found in anaerobic habitats all have mitochondria or mitochondria-like organelles called hydrogenosomes, and pyruvate oxidation via H_2-excretion appears to be central to their anaerobic lifestyle (Fenchel and Finlay 1991). Taxa known to inhabit anoxic marine habitats include the karyorelictids, prostomatids, haptorids, trichostomatids, entodiniomorphids, suctorids, scuticocilliatids, heterotrichids, odontostomatids, oligotrichids, and hypotrichids, some of which may be facultative anaerobes (Fenchel and Finlay 1995; Corliss 1979). Ciliates are often important grazers of abundant prokaryotes found along oxyclines, including in the deep sea (e.g., Lin et al. 2008; Pachiadaki et al. in press), and hence, may contribute significantly to nutrient cycling along OMZ oxygen gradients. *Perispira ovum*, a litostomatean ciliate, has been shown to prey heavily on euglenids, sequestering their chloroplasts and mitochondria as a potential mechanism to cope with low oxygen, thus acquiring the ability to function as mixotrophs in environments where mixotrophy confers a competitive advantage (Johnson et al. 1995). The lifestyle of litostomatean ciliates in the core of this OMZ, well outside of the photic zone, can only be speculated upon, and if these ciliates are active, other adaptive mechanisms may be present within such taxa not normally described from dark, anoxic, deep water columns. Within both of these diverse groups (ciliates and euglenids detected in this study), physiological adaptions to oxygen depletion and mutualistic associations with prokaryotes may enable persistence in OMZs.

Dinoflagellates of the order Syndiniales showed clear depth-specific differences. This group predominated within the 0.2-1.6 µm size fraction, increasing in richness at the OMZ core and contributing to an overall peak in diversity at this depth (Figure 2.1 and

2.2E). The Syndiniales are thought to be exclusive endoparasites of macrozoans and other protists (Moreira and Garcia, 2003, Chambouvet et al. 2008) and are frequently reported as the most abundant picoeukaryotes in environmental clone libraries from diverse ocean sites (Lopez-Garcia et al. 2001, Not et al. 2007, Countway et al. 2007, Sauvedat et al. 2010). Syndiniales are known to play important roles in controlling blooms of different taxa, and these parasitoids complete their life cycle in less than 3 days, releasing hundreds of free-living dinospores (<2μm) in the process (Chambouvet et al. 2008, Guillou et al. 2008). The abundance of dinospores may contribute to dominance of Syndiniales in many environmental surveys. Although the diversity and environmental distribution of Syndiniales are not fully understood, prior evidence shows that Groups I and II are present at hypoxic and suboxic sites (Guillou et al. 2008), with Group II more common than Group I below the photic zone. Here, Syndiniales Group II clades 5, 21, and 30 increased in relative abundance within the ETSP OMZ compared to more oxygenated waters above and below. Conversely, Clade 7 (Group II), the dominant Syndiniales group within the oxycline and in the oxic depths, decreased in abundance at the OMZ core. Some Group II clades (specifically 1,2, and 14) are parasites of other dinoflagellates and appear to show host specificity (Chambouvet et al. 2008). It is unclear whether the vertical distribution and ecological contributions of these and other Syndiniales clades in the ETSP OMZ are driven by free-living Syndiniales cells or by associations with hosts whose distributions vary with depth (assuming host lysis and release of cells into the picoeukaryotic fraction). Indeed, clone libraries from anoxic waters appear to contain members of the *Dinophysceae* sharing low similarity to known

dinoflagellates who may serve as hosts to Syndiniales (Stoeck et al. 2003, Wylezich et al. 2011).

Radiolarian protists related to *Acantharea* were also enriched in the OMZ. This may suggest they are also transported to the OMZ interior as sinking cells from the surface, potentially as aggregates or in association with fecal pellets (Turner, 2002). However, sequences affiliating with Radiolaria are frequently extremely abundant in environmental surveys of oxygen-depleted water columns (e.g., Edgcomb et al. 2011; Countway et al. 2010; Not et al. 2008; Table 2.2). In the ETSP, acantharids peaked in proportional abundance at the upper portion of the OMZ (110 m) and again at 320 m in the anoxic core. Acantharids are mixotrophs that are common in surface waters, especially in eutrophic environments (Gilg et al. 2010), where they have been shown to host large numbers of photosynthetic eukaryotic symbionts (Michaels 1991, Michaels et al. 1995). They are thought to be the most significant biological regulators of strontium in the ocean as they construct their cell walls entirely of strontium sulfate. Interestingly, these taxa are in very low abundance (<3%) at 200 m and at intermediate abundance (14-15%) in the oxic 32 m and 70 m samples. The acantharid-affiliated OTUs that were abundant in the OMZ at 110 and 320m were also present at surface depths, suggesting the transport of these taxa into the OMZ via sinking particles. However, we did not observe an even distribution of these signatures at depths between the surface and 320m, and so the patchy distribution of this group is inconsistent with a uniform flux of dead or dying radiolarians from the photic zone. Measurements of absolute abundance along with indicators of metabolic activity would be required to confirm and interpret this pattern.

Over half of all sequences at the surface (>1.6 μm fraction) were affiliated with metazoans (Figure 2.2A and 2.2D). While the dominant urochordate signal detected at the surface was lost in the transition into the OMZ, cnidarian, ctenophore, and *Maxillipoda* (copepod)- associated sequences detected at the surface were present in anoxic depths at low abundance (<10% of total sequences). Sequences matching jellyfish (Cnidaria) dominated the metazoan sequence pool in the anoxic zone. Diverse jellyfish have been observed in the oxycline and upper hypoxic layers of the OMZ of the Eastern Tropical North Pacific (Beatteay 2012, unpublished), consistent with the observed tolerance of many jellyfish to hypoxia (Elliot et al. 2012). Copepods have also been observed in diverse low oxygen waters where they exhibit species-specific depth distributions (Saltzman et al. 1997) and may possess physiological adaptations such as reduced respiration that allow persistence for brief periods under hypoxia (Teuber et al. 2013). Here, sequences matching copepods increased in relative abundance from the oxycline to the anoxic OMZ core before declining again at 1000 m, raising the potential that these organisms may occupy a niche within the anoxic zone where they may escape predation or have less competition. Additional taxonomic and physiological characterizations are needed to confirm the presence of a metazoan community adapted specifically to the extreme oxygen depletion observed in the ETSP.

2.5.1 Comparisons to other low oxygen sites

The taxonomic trends in the ETSP show broad similarities to those observed at other low oxygen marine sites, as well as notable differences, which may reflect the unique habitat characteristics of an OMZ. The relative dominance of dinoflagellate and radiolarian signatures in the ETSP OMZ has been observed at several other oxygen-

depleted sites in upwelling regions, including the Black Sea, Mariager Fjord, North Pacific, Saanich Inlet, and the Gulf of Mexico (Table 2.2), although due to potentially high ribosomal RNA gene copy numbers in these groups, these observations must be interpreted with caution until they are confirmed by microscopy or another method. Signatures of ciliates and euglenids were more abundant in our libraries from within the OMZ core and are also common in molecular surveys from other low oxygen environments. For example, litostomatean ciliates were identified in an Eastern Mediterranean anoxic basin (Edgcomb and Bernhard, 2013) and signatures of other diverse ciliates, including novel lineages have been recovered from Mariager fjord, the Black Sea, Baltic Sea, and the North Pacific (Table 2.2). Euglenid signatures were abundant in molecular surveys of both the Black Sea and Cariaco Basin (Table 2.2). Stramenopiles appear to be common across anoxic environments, but were rarely detected in our ETSP OMZ dataset.

The sequences recovered from many of these "low oxygen" groups vary by region and often do not match known sequences in public databases that would enable finer taxonomic assignments (Orsi et al. 2012, Stock et al. 2012, Edgcomb and Bernhard, 2013). This highlights a general gap in knowledge of the underlying diversity in oxygen-depleted waters. Furthermore, studies of eukaryote diversity in low-oxygen waters have utilized an array of sampling methods, sample sizes, and analysis tools, making it challenging to directly compare diversity metrics across sites (Stock et al. 2009, Edgcomb et al. 2011, Orsi et al. 2011, Rocke et al. 2012). For example, the analysis of multiple size fractions in our study identifies contrasting patterns of phylogenetic diversity with depth in the OMZ, suggesting knowledge of the effect of environmental

gradients (e.g, oxygen) on community structure may be biased depending on sample collection and processing strategy. Together, the increasing volume of sequence data from diverse low oxygen regions suggests geographic differences in protist communities in different oxygen-poor habitats, and high levels of uncharacterized microeukaryote diversity (Stoeck and Epstein, 2003, Zuendorf et al. 2006, Not et al. 2008, Orsi et al. 2012, Schnetzer et al. 2011). Resolving linkages between environmental parameters and diversity of microeukaryotes in OMZs will require application of uniform sampling methods across diverse low oxygen waters, as well as studies that couple gene surveys to genomic or ecophysiological measurements to link OTU distributions to functional activity.

2.6 Acknowledgements

We thank Jess Bryant, Adrian Sharma, Jaime Becker and the captain and crew of the R/V Melville for help in sample collection, and Tsultrim Palden for help in DNA library preparation and pyrosequencing. This work is a contribution of the Center for Microbial Oceanography: Research and Education (C-MORE) and was made possible by generous support from the National Science Foundation (1151698 to FJS and EF0424599 to EFD), the Gordon and Betty Moore Foundation (EFD), and the Agouron Institute (EFD). Edgcomb's involvement was supported by contributions from the Woods Hole Oceanographic Institution Director of Research and Ocean Life Institute.

2.7 References

Beatteay, SL. (2012). Impact of hypoxia on the vertical distribution of jellyfish in the Eastern Tropical Northern Pacific. University of Washington Senior Thesis.

Bernard C., and Fenchel T. (1996). Some microaerobic ciliates are facultative anaerobes. *Eur. J. Protistol*, 32, 293–297.

Bernhard JM, Buck KR, Farmer MA, Bowser SS. (2000). The Santa Barbara Basin is a symbiosis oasis. *Nature*, 403, 77–80.

Caporaso, J. G. et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7, 335–336.

Chambouvet A., Morin P., Marie D., Guillou L. (2008). Control of toxic marine dinoflagellate blooms by serial parasitic killers. *Science*, 322, 1254–1257.

Corliss, J.O. (1979). *The Ciliated Protozoa*. Pergamon Press, Oxford.

Countway P.D., Gast R.J., Dennett M.R., Savai P., Rose J.M. & Caron D.A. (2007). Distinct protistan assemblages characerize the euphotic zone and deep sea (2500 m) of the western North Atlantic (Sargasso Sea and Gulf Stream). *Environ Microbiol*, 9, 1219–1232.

Countway P.D., Vigil P.D., Schnetzer A., Moorthi S.D., Caron D.A. (2010). Seasonal analysis of protistan community structure and diversity at the USC Microbial Observatory (San Pedro Channel, North Pacific Ocean). *Limnol. Oceanogr*., 55, 2381–96.

Diez, B. Pedros-Alio, C., and Massan, R. (2001). Study of genetic diver-sity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Applied and Environmental Microbiology*, 67, 2932-2941.

Edgcomb V., Breglia S. A., Yubuki N., Beaudoin D., Patterson D. J., Leander B. S., Bernhard J. M. (2010). Identity of epibiotic bacteria on symbiontid euglenozoans in O2-depleted marine sediments: evidence for symbiont and host co-evolution. *ISME J.*, 5, 1–13.

Edgcomb V., et al. (2011). Protistan microbial observatory in the Cariaco Basin, Caribbean. I. Pyrosequencing vs Sanger insights into species richness. *The ISME Journal*, 5, 1344–1356.

Edgcomb, V.P., W. Orsi, H.-W. Breiner, A. Stock, S. Filker, M.M. Yakimov, and T. Stoeck. (2011). Novel kinetoplastids associated with hypersaline anoxic lakes in the Eastern Mediterranean deep-sea. *Deep-Sea Research I*, 58, 1040-1048.

Edgcomb, V., and Bernhard, J. (2013). Heterotrophic protists in hypersaline microbial mats and deep hypersaline basin water columns. *Life*, 3:2, 346-362.

Edgcomb, V.P. and Pachiadaki M. (2014). Ciliates along oxyclines of permanently stratified marine water columns. *J. Euk. Microbiol*, 4, 434-45.

Elliott, D. T., Pierson, J. and Roman, M. (2012). Relationship between environmental conditions and zooplankton community structure during summer hypoxia in the northern Gulf of Mexico. *J. Plankton Res.*, 34, 443-453.

Faith DP. (1992). Conservation evaluation and phylo- genetic diversity. *Biol Conserv*, 61, 1–10.

Fenchel, T., & Finlay, B.J. (1991). The biology of free-living anaerobic ciliates. *Europ J Protistol*, 26, 201-215.

Fenchel, T., & Finlay, B.J. (1995). *Ecology and Evolution in Anoxic Worlds*, Oxford Series in Ecology and Evolution, R.M. May and P.H. Harvey, Eds, Oxford University Press, New York, 276.

Fogg, G.E. (1995). Some comments on picoplankton and its importance in the pelagic ecosystem. *Aquatic Microbial Ecology*, 9, 33-39.

Ganesh, S., Parris, D.J., DeLong, E.F., Stewart, F.J. (2014). Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *ISME J.* 8, 187–211.

Gilg I.C., Amaral-Zettler L.A., Countway P.D., Moorthi S., Schnetzer A., et al. (2010). Phylogenetic affiliations of mesopelagic acantharia and acantharian-like environmental 18S rRNA genes off the southern California coast. Protist, 161, 197–211.

Grossart, H. P., Hietanen S. & Ploug H. (2003). Microbial dynamics on diatom aggregates in Øresund, Denmark. *Marine Ecology Progress Series*, 249, 69-78.

Guillou L., et al. (2008). Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environ Microbiol*, 10, 3349–3365.

Guillou L., Bachar D., Audic S., Bass D., Berney C., et al. (2012). The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Res*, 41, D597–D604.

Hayward B.H., Drosteb R., Epstein S.S. (2003). Interstitial ciliates: benthic microaerophiles or planktonic anaerobes? *J Eukaryot Microbiol*, 50, 356−359.

Johnson P.W., Donaghay P.L., Small E.B., Sieburth J.M. (1995). Ultrastructure and ecology of Perispira ovum (Ciliophora:Litostomatea): an aerobic, planktonic ciliate that sequesters the chloroplasts, mitochondria, and paramylon of Euglena proxima in a micro-oxic habitat. *Euk Microbiol*, 42, 323–335.

Kamp A., de Beer D., Nitsch JL., Lavik G., Stief P. (2011). Diatoms respire nitrate to survive dark and anoxic conditions. *Proc Natl Acad Sci*, 108, 14, 5649-54.

Lam, P. et al. (2009). Revising the nitrogen cycle in the Peruvian oxygen minimum zone. *Proceedings of the National Academy of Sciences* 106, 4752-4757.

Levin LA. (2003). Oxygen minimum zone benthos: Adaptation and community response to hypoxia. *Oceanography and Marine Biology: an Annual Review*, 41, 1–45.

Li XN., Taylor GT., Astor Y., Scranton MI. (2008). Sulfur speciation in the Cariaco Basin with reference to chemoautotrophic production. *Mar Chem.*, 112, 53–64.

Lin XJ., Scranton MI., Chistoserdov A., Varela R., Taylor GT. (2008). Spatiotemporal dynamics of bacterial populations in the anoxic Cariaco Basin. *Limnol Oceanogr*., 53, 37–51.

López-García P, Rodríguez-Valera F, Pedrós-Alió C, Moreira D (2001). Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature*, 409, 603–607.

Lozupone C, and Knight R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microb* 71, 8228–8235.

Lynn, D. (2008). *The ciliated protozoa: characterization, classification, and guide to the literature*, 3rd Edition, Springer, New York.

Massana R, Pernice M, Bunge JA, Del Campo J. (2011). Sequence diversity and novelty of natural assemblages of picoeukaryotes from the Indian Ocean. *ISME J*, 5, 184–195.

Michaels, AF. (1991). Acantharian abundance and symbiont productivity at the VERTEX seasonal station. *Journal of Plankton Research*, 13, 399-418.

Michaels AF, Caron DA, Swanberg NR, Howse FA, and Michaels CM. (1995). Planktonic sarcodines (acantharia, radiolaria, foraminifera) in surface waters near Bermuda: Abundance, biomass and vertical flux. *J Plankton Res*, 17, 131–163.

Montero, M., Daneri, G., Cuevas, L.A., González, H.E., Jacob, B., Lizárraga, L., and Menschel, E. (2007). Productivity cycles in the coastal upwelling area off Concepcion: the importance of diatoms and bacterioplankton in the organic carbon flux. *Prog Oceanogr*, 75, 518–530.

Moon-van der Staay SY, De Wachter R & Vaulot D. (2001). Oceanic 18S rDNA sequences from picoplankton reveal new eukaryotic lineages. *Nature*, 409, 607–610.

Moreira D., López-García P. (2003). Are hydrothermal vents oases for parasitic protists? *Trends Parasitol*, 19, 556–558.

Neira, C., Sellanes, J., Levin, L.A. & Arntz, W.E. (2001). Meiofaunal distributions on the Peru margin: relationship to oxygen and organic matter availability. *Deep-Sea Research* 48, 2453–2472.

Not, F., Gausling, R. Azam, F., Heidelberg, J.F., And Worden, A. (2007). Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environ. Microbiol*, 9, 1233–1252.

Not F, Latasa M, Scharek R, Viprey M, Karleskind P, et al. (2008). Phytoplankton diversity across the Indian Ocean: A focus on the picoplanktonic size fraction. *Deep - Sea Research Part I – Oceanographic Research Papers*, 55, 1456–1473.

Orsi W, Edgcomb V, Jeon S, Leslin C, Bunge J, Taylor GT et al. (2011). Protistan microbial observatory in the Cariaco Basin, Caribbean. II. Habitat specialization. *ISME J*, 5, 1357–1373.

Orsi W., Song Y. C., Hallam S., Edgcomb V. (2012). Effect of oxygen minimum zone formation on communities of marine protists. *ISME J*., 6, 1586–15601.

Orsi, W. and Edgcomb, V. (2013). Microbial eukaryotes in marine oxygen minimum zones. *Polyextremophiles*, 27, 485-497.

Pachiadaki, M., Taylor, C., Oikomomou, A., Yakimov, M., Stoeck, T., Edgcomb, V.P. (2014)**.** Grazing studies conducted *in situ* reveal protist turnover of bacterial biomass in the deep E. Mediterranean Sea. In press, *Deep Sea Res II*.

Revsbech NP, Larsen LH, Gundersen JK, Dalsgaard T, Ulloa O, Thamdrup B. (2009). Determination of ultra-low oxygen concentrations in oxygen minimum zones by the STOX sensor. *Limn Oceanogr: Methods*, 7, 371–381.

Rocke, E., Hongmei, J., and Liu, H. (2012). Phylogenetic composition and distribution of picoeukaryotes in the hypoxic northwestern coast of the Gulf of Mexico, *Microbiology Open*, 2, 1, 130-143.

Saltzman, J., Wishner, K.F., (1997). Zooplankton ecology in the eastern Pacific oxygen minimum zone above a seamount: 2. Vertical distribution of copepods. *Deep-Sea Research I*, 44, 931–954.

Sauvadet A-L., Gobet A., Guillou L. (2010). Comparative analysis between protist communities from the deep-sea pelagic ecosystem and specific deep-hydrothermal habitats. *Env. Microbiol* 12, 2946–2964.
Schnetzer A., et al (2011). Depth matters: microbial eukaryote diversity and community structure in the eastern North Pacific revealed through environmental gene libraries. *Deep Sea Res. Part 1 Oceanogr. Res. Pap*, 58, 16–26.
Smetacek, V.S. (1985). Role of sinking in diatom life-history cycles: Ecological, evolutionary, and geological significance, *Marine Biology*, 84, 239-251.
Stevens H., and Ulloa O. (2008). Bacterial diversity in the oxygen minimum zone of the eastern tropical South Pacific. *Environmental Microbiology*, 10, 1244–1259.
Stock A, Bunge J, Jürgens K & Stoeck T (2009). Protistan diversity in the suboxic and anoxic waters of the Gotland Deep (Baltic Sea) as revealed by 18S rRNA clone libraries. *Aquat Microb Ecol*, 55, 267–284.
Stock A, Breiner HW, Pachiadaki M, Edgcomb V, Filker S, et al. (2012). Microbial eukaryote life in the new hypersaline deep-sea basin Thetis. *Extremophiles*, 16, 21–34.
Stoeck T., Taylor GT., & Epstein SS. (2003). Novel eukaryotes from a permanently anoxic Cariaco Basin (Caribbean Sea). *Appl Environ Microb*, 69, 5656–5663.
Tang KW., Turk V., Grossart HP. (2010). Linkage between crustacean zooplankton and aquatic bacteria. *Aquat Microb Ecol*, 61, 261-277.
Taylor, G.T., (1982). The role of pelagic heterotrophic protozoa in nutrient cycling: a review. *Annales Institute Oceanographie, Paris*, 58, 227-241.
Taylor GT., Scranton MI., Iabichella M., Ho T., Thunell RC., Muller-Karger F., et al. (2001). Chemoautotrophy in the redox transition zone of the Cariaco Basin: a significant midwater source of organic carbon production. *Limnol Oceanogr.*, 46, 148–163.
Teuber L., Kiko R., Séguin F., Auel H. (2013). Respiration rates of tropical Atlantic copepods in relation to the oxygen minimum zone. *Journal of Experimental Marine Biology and Ecology*, 448, 2836.
Thamdrup B., Dalsgaard T., Jensen MM., Ulloa O., Farías L., Escribano R. (2006). *Limnol Oceanogr*, 51, 2145–2156.
Thamdrup B., Dalsgaard T., and Revsbech NP. (2012). Widespread functional anoxia in the oxygen minimum zone of the eastern South Pacific. *Deep Sea Res*, 65, 36–45.
Turner J.T. (2002). Zooplankton fecal pellets, marine snow and sinking phytoplankton blooms. *Aquat. Microb. Ecol.*, 27, 57–102.
Ulloa O., Canfield D. E., Delong E. F., Letelier R. M., Stewart F. J. (2012). Microbial oceanography of anoxic oxygen minimum zones. Proc. *Natl. Acad. Sci*., 109, 15996–16003.
Worden, A.Z., Nolan, J.K., Palenik, B. (2004). Assessing the dynamics and ecology of marine picophytoplankton: The importance of the eukaryotic component. *Limnol. Oceanogr*. 49, 168–179.
Wright JJ., Konwar KM., Hallam SJ. (2012). Microbial ecology of expanding oxygen minimum zones. *Nat. Rev. Microbiol*, 10, 381–394.
Wylezich C., Jürgens K. (2011). Protist diversity in suboxic and sulfidic waters of the Black Sea. *Environ Microbiol*, 13, 2939–2956.
Yubuki N, Edgcomb VP, Bernhard JM, Leander BS. (2009). Ultrastructure and molecular

phylogeny of Calkinsia aureus: cellular identity of a novel clade of deep-sea euglenozoans with epibiotic bacteria. *BMC Microbiol*, 9, p.16.

Zuendorf A, Bunge J, Behnke A, Barger KJA, Stoeck T. (2006). Diversity estimates of microeukaryotes below the chemocline of the anoxic Mariager Fjord, Denmark. *Fems Microbiol Ecol*, 58, 476–491.

CHAPTER 3.

WHOLE GUT MICROBIOME COMPOSITION OF DAMSELFISH AND CARDINALFISH BEFORE AND AFTER REEF SETTLEMENT

Disclaimer: This chapter was published with the same title, along with the supplementary material in Appendix B, in the journal PeerJ on August 31, 2016

Citation: Parris, D. J., Brooker, R. M., Morgan, M. A., Dixson, D. L., & Stewart, F. J. (2016). Whole gut microbiome composition of damselfish and cardinalfish before and after reef settlement. *PeerJ*, *4*, e2412.

3.1 Abstract

The Pomacentridae (damselfish) and Apogonidae (cardinalfish) are among the most common fish families on coral reefs and in the aquarium trade. Members of both families undergo a pelagic larvae phase prior to settlement on the reef, where adults play key roles in benthic habitat structuring and trophic interactions. Fish-associated microbial communities (microbiomes) significantly influence fish health and ecology, yet little is known of how microbiomes change with life stage. We quantified the taxonomic (16S rRNA gene) composition of whole gut microbiomes from ten species of damselfish and two species of cardinalfish from Lizard Island, Australia, focusing specifically on comparisons between pelagic larvae prior to settlement on the reef versus post-settlement juvenile and adult individuals. On average, microbiome phylogenetic diversity increased from pre- to post-settlement, and was unrelated to the microbial composition in the surrounding water column. However, this trend varied among species, suggesting stochasticity in fish microbiome assembly. Pre-settlement fish were enriched with bacteria of the Endozoicomonaceae, Shewanellaceae, and Fusobacteriaceae, whereas settled fish harbored higher abundances of Vibrionaceae and Pasteurellaceae. Several individual operational taxonomic units, including ones related to *Vibrio harveyi*, *Shewanella sp.*, and uncultured *Endozoicomonas* bacteria, were shared between both pre and post-settlement stages and may be of central importance in the intestinal niche across development. Richness of the core microbiome shared among pre-settlement fish was comparable to that of settled individuals, suggesting that changes in diversity with adulthood are due to the acquisition or loss of host-specific microbes. These results

identify a key transition in microbiome structure across host life stage, suggesting changes in the functional contribution of microbiomes over development in two ecologically dominant reef fish families.

3.2 Introduction

Animals harbor diverse microbial communities that influence key aspects of host health, development, and behavior. In most vertebrates, the majority of microorganisms inhabit the gastrointestinal (GI) tract at an abundance of potentially trillions of cells whose collective genome may be orders of magnitude larger than that of the host (Whitman et al. 1998; Backhed et al. 2005; Gill et al. 2006). A wealth of studies in both fish and mammalian systems now confirm the importance of a gut microbiome to host health, fitness, and development (Rawls et al. 2004, Bates et al. 2006, Chung et al. 2012; Lee & Hase 2014), with alterations of microbiome composition affecting such properties as host immunity, susceptibility to toxins, adiposity, efficiency of nutrient acquisition, behavior and mood, and chemical signaling among conspecifics (Turnbaugh et al. 2006; Round & Mazmanian 2009; Heijtz et al. 2011; Greenblum et al. 2012; Theis et al. 2013; Zheng et al. 2013, Wada-Katsumata et al. 2015).

Despite recognition of the important role of microbiomes in host ecology (Wong et al. 2012), microbiomes remain unexplored for the vast majority of ecologically relevant taxa, including the most diverse of the vertebrate groups, the teleost fishes. Fish account for almost half of all vertebrate species on earth and span a wide spectrum of habitats, host ranges, physiologies, and ecological strategies. Studies of fish gut microbiomes have focused primarily on commercial or model species, with most targeting only a small number of host or microbial taxa (Roeselers et al. 2011; Clements

et al. 2014; Llewellyn et al. 2014). These studies raise the possibility of a core set of fish gut microbes shared across diverse hosts, notably a dominance by bacteria of the Gammaproteobacteria and Firmicutes, including unique fish-associated strains of the Vibrionales and Clostridia (Spanggaard et al. 2000; Al-Harbi & Naim Uddin 2004; Martin-Antonio et al. 2007; Sullam et al. 2012; Xing et al. 2013; Llewellyn et al. 2014). Significant variation in the fish gut microbiome has also been reported, with changes in taxonomic composition shown to affect host immunity, nutrient acquisition, and epithelial differentiation (Rawls et al. 2004; Bates et al. 2006, 2007; Cheesman & Guillemin 2007; Cheesman et al. 2010; Kanther & Rawls 2010, Ghanbari et al. 2015). Fish microbiome composition has been linked to diverse factors including host type (Ye et al. 2014; Givens et al. 2015, Hennersdorf et al. 2016), trophic ecology and diet (Bolnick et al. 2014a,b; Miyake et al. 2015; Sullam et al. 2015), and environmental conditions (e.g., salinity; Sullam et al. 2012; Schmidt et al. 2015).

The composition of the fish microbiome can also change significantly over an individual's lifespan (Bakke et al. 2015; Burns et al. 2015, Llewellyn et al. 2015). However, the timing, determinants, and magnitude of these shifts are unclear for most taxa, and likely highly variable among species with differing life histories. On coral reefs, for example, >90% of fish species undergo a dispersing or planktonic stage as either larvae or juveniles, followed by an adult phase marked by settlement into a constrained territory on the reef (Leis et al. 1991). For such species, microbiome structure is hypothesized to be shaped by a combination of factors that co-associate with host life stage, including diet shifts, immune system and GI tract development, and

differences in the composition of the external microbial community between pre- and post-settlement habitats.

Comparing reef fish microbiomes over developmental stages will help identify shifts in the contribution of the microbiome to host health and behavior, thereby helping reach a holistic understanding of reef ecology. Collectively, coral reefs harbor over 2500 fish species engaged in a complex web of interactions, including predation, herbivory, corallivory, and symbiosis (Bellwood et al. 2004; Allen 2014). These interactions together modulate material and energy transport and ecosystem structuring on coral reefs, with important consequences for reef-building corals that rely on fish waste for nutrients, or fish herbivory to limit competition with benthic macroalgae (Mumby et al. 2006; Burkepile & Hay 2008; McCauley et al. 2010). Such interactions may also be important vectors for moving microbes between reef habitats. A handful of studies have characterized gut microbiomes of coral reef fish, identifying a role for both diet type and host identity in shaping microbiome composition (Smriga et al. 2010, Miyaki et al. 2015). However, the microbiomes of the vast majority of reef fishes, including some of the most abundant and ecologically relevant taxa, remain uncharacterized.

Damselfish (Pomacentridae) and cardinalfish (Apogonidae) are among the most diverse and abundant families of reef fishes and exhibit similar ecological strategies. Species within these families occupy critical positions in reef food webs as generalist planktivorous consumers, subsisting on a variety of algae, plankton, and benthic invertebrates (Emery et al. 1973, Emery et al. 1980, Marnane and Bellwood, 2002). These fish also constitute major fractions of the diet of larger piscivorous reef fish (Beukers-Stewart et al. 2004). Most species of damselfish and cardinalfish spend their

first 12 to 35 days of life as members of the off-reef plankton community before settling back onto the reef as juveniles (Wellington and Victor, 1989, Leis et al. 1991, Leis et al. 1996). Post-settlement, damselfishes (including members of the *Chromis*, *Pomacentrus*, and *Dascyllus* genera) occupy distinct microhabitats using various types of coral for habitat and cover (Wilson et al. 2008). Many Pomacentrids are highly territorial and influence turf algae composition through grazing (Hinds, 1987, Klumpp and Polunin, 1989). Apogonids also occupy coral microhabitats during the day when they are mostly inactive but leave coral cover at night to feed throughout the water column (Gardiner and Jones, 2005). Fish species of the Apogonidae often dominate nocturnal planktivore assemblages (Manane and Bellwood, 2002). In this study, we quantified microbiome taxonomic composition (16S rRNA gene diversity) in whole guts from replicate individuals from ten species of damselfish and two species of cardinalfish (Table S1). We focused on comparisons between pre-settlement planktonic larvae and post-settlement individuals (juveniles through to adults) to determine whether microbiome composition varied between ontogenetic stages and, if so, whether this variation was related to differences in the microbial composition of the surrounding water column.

3.3 Materials and Methods

3.3.1 Sample collection

Fishes were collected over a 3-4 day period from the fringing reefs surrounding Lizard Island, Australia (14°40.08'S 145°27.34'E) in February 2014. Water depths at collection sites were approximately 5-10 meters. Larval, pre-settlement individuals (8-16 mm total length, Table B.1) were defined as samples collected from floating light traps deployed overnight at a single site. These traps sampled the upper meter of the water

column above the reef. Post-settlement stage juveniles/adults (16-36 mm total length, Table B.1) were individuals hand-collected on SCUBA at depth from nearby reefs using dip nets. Post-settlement fish in this study were likely sexually immature (juveniles) and recently settled, given their small size (sexual maturity in most Pomacentrids typically occurs at lengths >60 mm, Schmale et al. 1986, Kavanaugh, 2000). Net capture of settled fish followed mild sedation with a clove oil/ethanol mixture dispensed via spray bottle. All fish were euthanized by immersion in a seawater-ice slurry combined with an overdose of clove oil and stored in ethanol for further analysis. Water column samples (n=11) were collected on SCUBA at the same depth and location of post-settlement fish collection sites and the light trap site using sterile plastic 1L bottles. Bottles were first opened within the water column to ensure collection at the desired location and prevent contamination. Microbial biomass in the sampled water was concentrated onto 0.2 μm Polyvinylidene Flouride (PVDF) filters through filtration. Filters were immediately placed in ethanol in cryovials and stored at room temperature until processing. Fish were identified to the lowest possible taxonomic unit, typically either genus or species, using Indo-Pacific reef fish field guides based on morphology. In total, 49 pre-settlement individuals were sampled (Table B.1), including individuals of the damselfish *Pomacentrus moluccensis* (n=4), *P. chrysurus* (n=4), *P. nagasakiensis* (n=4), *P. amboinensis* (n=4), *P. wardii* (n=4), *P. bankanensis* (n=4), *P. coelestis* (n=4), *Dascyllus aruanus* (n=3) and an unidentified Chromis species (n=7). Two species of pre-settlement stage cardinalfish, *Ostorhinchus doederleini* (n=4) and an unidentified Apogon species (n=7), were also collected. A total of 24 post-settlement individuals were sampled (Table S1), including individuals of 5 species of damselfish (*P. moluccensis* (n=6), *P.*

chrysurus (n=3), *A. polyacanthus* (n=3), *Dascyllus aruanus* (n=5) and an unidentified Chromis sp. (n=3)) and 1 cardinalfish (*O. doederleini* (n=4)). Note that the individuals of the unidentified *Chromis* species (collected both pre and post-settlement) were almost certainly either *Chromis viridis* or *Chromis atripectoralis*, both of which are similar ecologically and closely related. All research was reviewed and conducted under the guidelines of Georgia Tech IACUC #A14063 and Great Barrier Reef Marine Park Authority collection permits G13.36166.1 and animal ethics permits A1920. These methods were approved by Georgia Tech IACUC and Great Barrier Reef Marine Park Authority for use in this study.

3.3.2 DNA extractions

Whole fish samples were removed from storage containers and rinsed thoroughly with fresh, filter-sterilized ethanol to remove (potentially) surface-attached microbes. As dissection of only the intestinal contents was not possible due to the small size of most pre-settlement individuals, whole gut contents (stomach + intestines) were excised by dissection with a sterile razor from each individual. Excised contents were placed in a sterile 1.5 ml centrifuge tube and frozen until extraction. Bulk DNA was extracted from gut contents using the Qiagen DNeasy blood and tissue kit according to manufacturer instructions. To process water samples, each collection filter was transferred from ethanol into a clean, 1.5 ml centrifuge tube. The ethanol remaining in the original storage tube was vacuum-filtered onto a new 0.2 µm PVDF filter, which was then dried at room temperature and pooled with the initial collection filter. DNA was extracted from filters using a phenol:chloroform method. Briefly, cells were lysed by adding lysozyme (2 mg in 40 µl of lysis buffer per sample) directly to the pooled filters, capping, and incubating

for 45 min at 37ºC. Proteinase K (1 mg in 100 µl lysis buffer, with 100 µl 20% SDS) was added and samples incubated for an additional 2 hours at 55ºC. The lysate was removed, and DNA was extracted once with phenol:chloroform:isoamyl alcohol (25:24:1) and once with chloroform:isoamyl alcohol (24:1). Extracted DNA was concentrated by spin dialysis using Ultra-4 (100 kDa, Amicon) centrifugal filters and stored frozen.

3.3.3 PCR amplification and sequencing

High-throughput sequencing of dual-indexed PCR amplicons spanning the V3-V4 hypervariable regions of the 16S rRNA gene was used to assess gut microbiome taxonomic composition. Amplicons were synthesized using Platinum® PCR SuperMix (Life Technologies) with primers F515 (5'-GTGCCAGCMGCCGCGGTAA-3') and R806 (5'-GGACTACHVGGGTWTCTAAT-3', Caporaso et al., 2011), with both primers modified to include sample-specific barcodes and Illumina sequencing adapters according to Kozich et al. (2013). Ten nanograms of starting DNA was used as template for each PCR reaction. Amplification was performed using denaturation at 94°C (3 min), followed by 30 cycles of denaturation at 94°C (45 sec), primer annealing at 55°C (45 sec), primer extension at 72°C (90 sec), and a final extension at 72°C for 10 min. Amplicon products were verified using gel electrophoresis, purified using Diffinity RapidTip2 PCR purification tips (Diffinity Genomics, NY), and quantitated fluorometrically using the Qubit (Life Technologies). Amplicons from different samples were pooled at equimolar concentrations and sequenced on an Illumina MiSeq using a 500 cycle kit with 30% PhiX added to increase sequence diversity. All raw sequences are available in the NCBI Sequence Read Archive under BioProject ID PRJNA290348.

3.3.4 Sequence Analysis

Barcoded sequences were de-multiplexed and trimmed (length cutoff: 100 bp) and filtered to remove low quality reads (cutoff: Phred score of 25, averaged over all bases) using Trim Galore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Paired-end reads were then merged using FLASH (Magoč & Salzberg 2011), with an average read length of 250, and a fragment length of 300. Merged reads were analyzed using the QIIME pipeline (Caporaso et al. 2010). In QIIME, chimeric sequences were identified using USEARCH (Edgar, 2010) and removed from the input dataset. Merged, non-chimeric sequences were clustered into Operational Taxonomic Units (OTUs) at 97% sequence similarity using open-reference picking with the UCLUST algorithm (Edgar 2010). Singleton OTUs (those represented by only one sequence read in the combined dataset) were also removed at this step. Taxonomy was assigned to each OTU by comparison with the Greengenes database 13_5 release (DeSantis et al. 2006). Cyanobacterial and chloroplast sequences were manually removed from the OTU tables as these sequences can be assumed to reflect non-host-associated taxa, potentially those brought into the gut through food passage. Selected OTUs were queried via BLASTN against the NCBI-nr database to identify closest relatives.

OTU counts were rarefied (10 iterations) to 4262 sequences per sample and the rarified OTU table was used in downstream analysis. A Bonferroni-corrected ANOVA was used to identify microbial families significantly enriched in pre vs post-settlement individuals, with family composition data expressed as percentage of total reads. Alpha and beta diversity statistics for pre- and post-settlement individuals were calculated in QIIME using the core_diversity_analyes.py command. The unweighted Unifrac metric

was used to conduct principal coordinates analysis to visualize beta diversity. Significant differences in beta diversity were identified using a two-sided t-test (Bonferroni-corrected). Core microbiome analysis was performed in QIIME using the compute_core_microbiome.py script to identify individual OTUs shared by at least 70% of all individuals per sample grouping (all pre-settlement individuals vs. all post-settlement individuals). This script was also used to estimate the average number of OTUs shared between random pairs of individuals from each life stage, based on 100 random iterations per life stage. Indicator analysis was performed with the indicspecies package in R to identify genera (indicator taxa) significantly enriched according to life stage, taking into consideration both the extent to which an OTU is exclusive to a life stage (A component), as well as the relative frequency at which the OTU occurs in all individuals within a stage (B component)(Dufresne and Legendre 1997; Fortunato et al. 2013). Only those taxa occurring in more than half of all samples (B component >0.5) and for which permutation testing yielded p-values <0.05 are included in the output. Indicator values range from 0 to 1 with higher values reflecting stronger relationships between indicator taxa and the tested sample groupings (pre vs. post-settlement).

3.4 Results and discussion

3.4.1 General microbiome characteristics

A total of 49 pre-settlement and 24 post-settlement individuals representing ten Pomacentridae species and two Apogonidae species were collected from Lizard Island, Australia. Of these, five species - the pomacentrids *Pomacentrus moluccensis*, *P. chrysurus*, *Dascillus aruanus*, and an unidentified *Chromis* sp. (either *C. viridis* or *C. atripectoralis*) and the apogonid *Ostorhinchus doederleini* - were represented by both

pre- and post-settlement individuals, with the remaining species represented only in the pre-settlement sample set (n = 6 species) or only in the post-settlement set (n = 1).

A total of 3,503,605 16S rRNA gene sequences were obtained after quality control filtering (average: 41,709 per sample; range: 4262-238,858). After rarefaction, OTU counts per sample were highly variable, from 44 to 557 OTUs, with averages of 163 (±59 standard deviation (SD)) and 245 (±102 SD) for pre- and post-settlement stages, respectively. As observed in other fish species (Spanggaard et al. 2000; Al-Harbi & Naim Uddin 2004; Martin-Antonio et al. 2007; Sullam et al. 2012; Xing et al. 2013), Pomacentrid and Apogonid gut microbiomes were dominated by Gammaproteobacteria of the Pseudoaltermonadaceae, Endozoicimonaceae, Vibrionaceae, and Shewanellaceae, with these Families constituting 80% (±10.3 SD) and 67% (±6.7 SD) of all sequence reads in pre- and post-settlement fishes, respectively (Figure B.1), although key differences in the relative abundance of microbial families were also evident (discussed below, Table 3.1).

Table 3.1 Median % abundance of bacterial Families* in pre- versus post-settlement damselfish and cardinalfish (pooled).

	Pre		Post		
Family	med	MAD	med	MAD	Fold
Vibrionaceae	18.12	14.20	41.91	28.69	2.31
Endozoicimonaceae**	11.02	9.50	0.14	0.14	76.40
Pseudoalteromonadaceae	1.28	1.15	0.45	0.42	2.83
Rhodobacteraceae	1.24	1.01	1.35	1.22	1.09
Flavobacteriaceae	1.07	0.89	0.61	0.60	1.75
Alteromonadaceae	0.71	0.63	0.85	0.75	1.20
Shewanellaceae	0.68	0.65	0.11	0.11	5.99
Oceanospirillaceae	0.56	0.51	0.15	0.15	3.74
Pseudomonadaceae	0.29	0.24	0.49	0.45	1.70
Moraxellaceae	0.26	0.23	0.02	0.02	10.31
Pirellulaceae**	0.01	0.01	0.83	0.81	81.97
Chromatiales, Unknown	0.10	0.08	0.21	0.20	2.03

*only includes Families with median abundance > 0.2% in either pre or post datasets
**significant change pre vs. post ($P < 0.05$, ANOVA, Bonferroni corrected)
MAD, median absolute deviation
Fold, fold increase in median (pre to post); underlined indicates post to pre

It is important to note that our data do not differentiate communities based on location in the GI tract, as the small size of individuals prohibited separation of intestine from stomach contents. Thus, we cannot distinguish between resident gut microbiome members, which might be relatively more common in the intestine, versus members derived from food contents or passively from seawater intake, which might be relatively enriched in the stomach. Both host-adapted as well as transient microbiome members may nonetheless exert effects on host health, as either symbionts or potentially as inocula for resident populations. However, gut microbiome composition (weighted Unifrac; Pomacentridae + Apogonidae combined) of both pre- and post-settlement stages differed significantly from that of the surrounding seawater ($p=0.028$; Figure 3.1), which was instead enriched in Gammaproteobacteria of the Halomonadaceae family, Alphaproteobacteria, and Bacteroidetes (Figure B.2). Moreover, microbiome composition did not differ significantly between water collection sites (Figure B.3), suggesting fine-scale variation in local environments alone should not influence microbiome composition and an overall minimal effect of source water on gut microbiome structure.

3.4.2 Microbiome diversity and composition

Microbiomes showed a general trend toward greater complexity following settlement, although this trend was variable among species. The phylogenetic composition of the combined (Pomacentrid + Apogonid) microbiomes changed significantly from pre- to post-settlement groups (weighted Unifrac metric; $p=0.028$), with pre-settlement microbiomes showing less intra-individual variation compared to

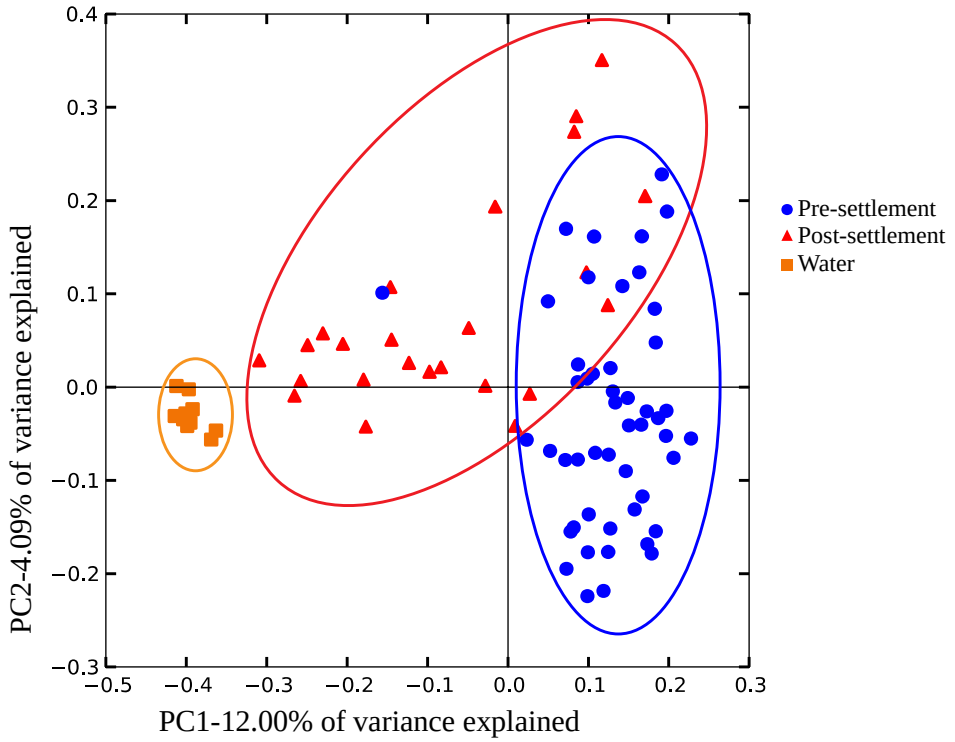

Figure 3.1. **Compositional relatedness of microbiome samples.** Principal coordinate analysis based on unweighted Unifrac distance with sequence data rarified to 4262 sequences per sample (49 pre-settlement, 24 post-settlement, 11 seawater samples). According to a two-sided t-test (Bonferroni-corrected), pre-settlement and post-settlement individuals are significantly different from one another ($p<0.05$) and both are significantly different from the water column ($p<0.05$) based on unweighted Unifrac distance.

those of post-settlement fish (Figure 1) and overall lower phylogenetic diversity (PD) (average: 20.3 ± 7 SD, versus 28.7 ± 12 SD post-settlement; p=0.001, t-stat = 3.5; Figure 3.2). Both pre- and post-settlement microbiomes exhibited substantially lower PD than the surrounding water column community (PD: 79.4 ± 23.5 SD; p=0.001). Of the five

species for which both pre- and post-settlement samples were available, three - *Chromis sp., O. doederleini,* and *D. aruanus* - exhibited significantly higher PD post-settlement (Figure 3.3, p=0.001). Greater microbiome diversity with age has been observed in other fish species, with suggested linkages to increases in diet complexity (Bolnick et al. 2014a,b; Miyake et al. 2015; Sullam et al. 2015), or potentially to differentiation of the GI tract into distinct niches with development. In contrast, elevated microbiome diversity in early life was observed in *P. chrysurus* and *P. moluccensis* (Figure 3.3, p=0.001), consistent with data from lab-reared zebrafish (Stephens et al 2015). This pattern suggests that pre-settlement microbiomes in these species are influenced primarily by transient associations with bacteria from the exterior environment or by higher diet complexity at the planktonic stage. As an individual matures, its microbiome may become more specialized in response to diverse factors, including diet shifts, immune system development, physiological and chemical changes along the GI tract, or decreased connectivity with the external environment. Indeed, species-specific microhabitat usage (i.e., selection of distinct coral types by different species) has been observed in some of the hosts examined here, including *P. moluccensis* and *P. nagasakiensis*. However, all settled hosts in our dataset are assumed to be generalist planktivores with fairly similar post-settlement diets based on prior work (Emery et al. 1973, Emery et al. 1980, Marnane and Bellwood, 2002). Although this study does not allow us to quantify the influence of

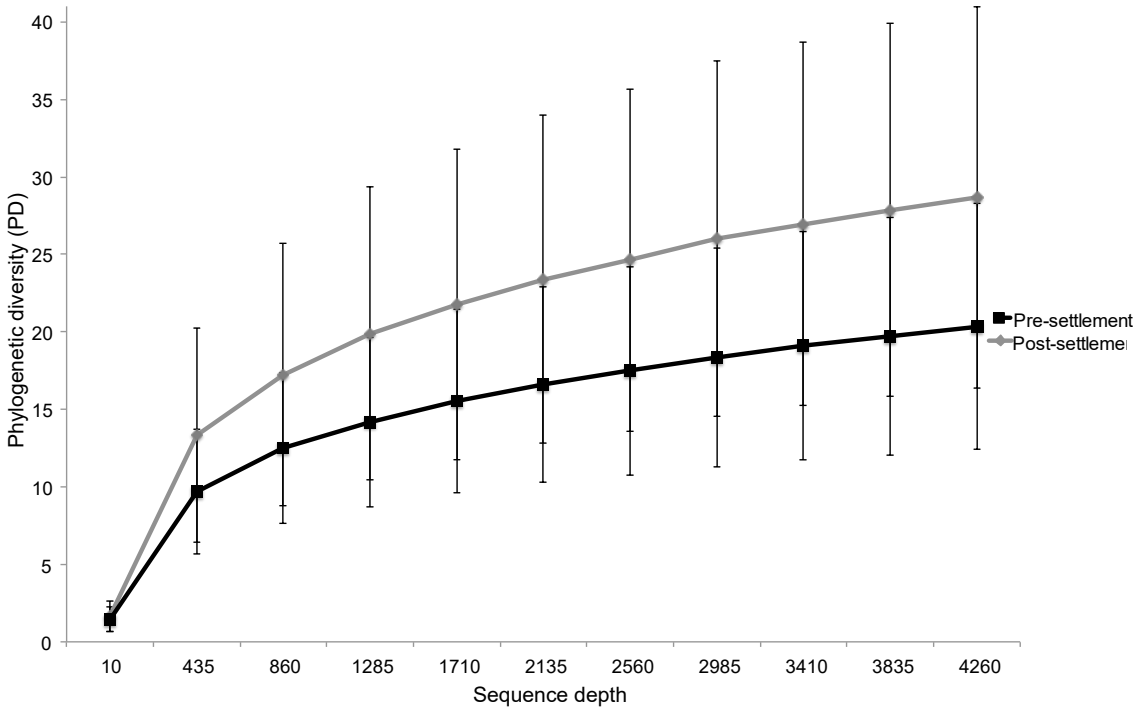

Figure 3.2. **Microbiome phylogenetic diversity (PD) as a function of sequence depth in pre- and post-settlement fish** (all species combined, based on data rarified to 4262 sequences per sample). Error bars are ± 1 standard error.

each of the factors discussed above on developmental shifts in microbiome composition, our data, combined with results from prior studies, suggest that microbiome diversity is dynamic over life stages and assembly processes vary on a per-species basis.

Key microbial taxa differed in abundance between pre- and post-settlement microbiomes (Table 3.1), although microbiome composition was relatively consistent within hosts per age group (Figure B.3). At the microbial family level, pre-settlement microbiomes were significantly enriched (>70-fold) in bacteria of the Endozoicimonaceae, while settled fish harbored over 2-fold higher median abundances of the abundant Vibrionaceae and an 80-fold enrichment of planctomycetes of the

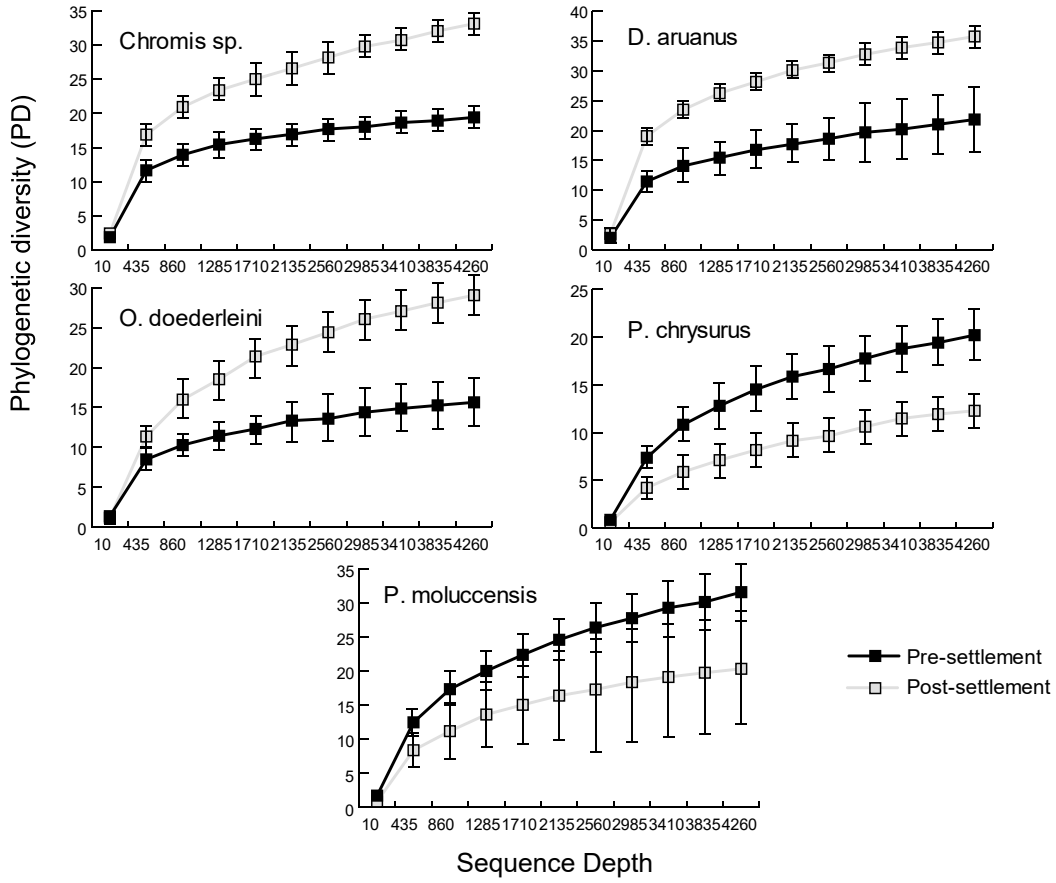

Figure 3.3. **Microbiome phylogenetic diversity (PD) as a function of sequence depth in five species** represented in both pre- and post-settlement datasets. Error bars are ± 1 standard error.

Pirellulaceae (Table 3.1). Differences in the relative abundance of microbes at the family level did not appear to be driven by a single host. For example, all pre-settlement individuals were enriched in bacteria of the Endozoicomonaceae compared to post-settlement hosts (Figure B.4). At the microbial genus level, both life stages were marked by sets of indicator taxa (Table 3.2). Indicator taxa represent individual OTUs showing significant enrichment (as a percent of total sequence reads) in one life stage vs. the other according to permutation testing ($p<0.05$, Dufresne and Legendre 1997; Fortunato et al. 2013). The number of indicators was considerably higher for settled fish (38) compared to larvae (5; Table 3.2), consistent with the increase in diversity associated with post-

settlement, perhaps reflective of an overall expansion in microbial niche breadth as discussed above. The marine bacterial genera *Kordia* (Flavobacteriia) and *Halomonas* (Gammaproteobacteria, Oceanospirillales) were the strongest indicators of pre-settlement stage (Table 3.2), followed by members of the *Arcobacter* (Epsilonproteobacteria, Campylobacterales), *Oceanospirillum* (Gammaproteobacteria, Oceanospirillales), and *Idiomarina* (Gammaproteobacteria, Alteromonadales), all taxa common to the marine environment but not often observed in fish microbiomes. Pre-settlement enrichment implicates these taxa as important during early development, more successful at colonizing and growing in young fishes, or commonly associated with pre-settlement food sources. Indicators of post-settlement spanned a wide phylogenetic breadth, with a number classifiable only to the order or family level, and the strongest indicators including diverse members of the Gammaproteobacteria (Vibrionaceae, Portiera), Firmicutes (Epulopiscium), Alphaproteobacteria (Kiloniellales, Hyphomicrobiaceae), Verrucomicrobia (Coraliomargarita), and Planctomycetes (Pirellulaceae) (Table 3.2). Enrichment of these taxa in fully developed, settled fish suggests the host GI tract as a specific niche for members of these groups.

Table 3.2. Indicator genera associated with pre- and post-settlement damselfish and cardinalfish (pooled).

Indicator genera	Indicator for pre or post settlement fish	Indicator value	p-value
Kordia	Pre	0.943	0.001
Halomonas	Pre	0.896	0.001
Arcobacter	Pre	0.869	0.001
Oceanospirillum	Pre	0.823	0.001
Idiomarina	Pre	0.751	0.002
unclassified Vibrionaceae	Post	0.997	0.01
unclassified Pirellulaceae	Post	0.922	0.001
unclassified Kiloniellales	Post	0.904	0.001
unclassified Gammaproteobacteria	Post	0.9	0.001
Coraliomargarita	Post	0.863	0.001
Portiera	Post	0.857	0.001
Epulopiscium	Post	0.857	0.001
unclassified Hyphomicrobiaceae	Post	0.851	0.001
Verrucomicrobium	Post	0.816	0.001
unclassified Rhizobiales	Post	0.786	0.001
Rhodospirillaceae	Post	0.779	0.001
unclassified Acidimicrobiales OCS155	Post	0.778	0.001
unclassified Oceanospirillales	Post	0.771	0.001
unclassified Peptostreptococcaceae	Post	0.771	0.001
unclassified Myxococcales	Post	0.768	0.001
unclassified Altermonadales DM60	Post	0.763	0.013
unclassified Phyllobacteriaceae	Post	0.761	0.001
Ferrimonas	Post	0.75	0.001
unclassified Thiohalorhabdales	Post	0.746	0.001
unclassified Flavobacteriales	Post	0.743	0.001
unclassified Cryomorphaceae	Post	0.733	0.038
unclassified Altermonadales	Post	0.733	0.038
unclassified Myxococcales DM27	Post	0.728	0.001
Balneola	Post	0.724	0.001
Crocinitomix	Post	0.723	0.001
unclassified Flammeovirgaceae	Post	0.722	0.002
unclassified Altermonadaceae	Post	0.722	0.035
unclassified Francisellaceae	Post	0.712	0.001
Saprospria	Post	0.692	0.003
Clostridium	Post	0.692	0.016
unclassified Rickettsiales	Post	0.688	0.001
Turicibacter	Post	0.682	0.001
unclassified Deltaproteobacteria GMD14H09	Post	0.678	0.006
unclassified Pasteurellaceae	Post	0.676	0.002
unclassified Actinomycetales	Post	0.675	0.001
Altermonadaceae HTCC2207	Post	0.67	0.001
unclassified Comamonadaceae	Post	0.664	0.004
Flavobacterium	Post	0.656	0.003

3.4.3 Comparison of the core microbiome between pre and post-settlement individuals

Although egg-tending behavior is common in many pomacentrids, fish are presumed sterile at birth and fish gut microbiome is assumed to be inoculated primarily from the environment and via food items early in life (Nayak 2010). If the pre-settlement, pelagic diet is similar across species, it is reasonable to predict that larval

individuals of different fish species will share a greater number of microbial taxa (i.e., larger core microbiome). The core microbiome might then be expected to shrink post-settlement as species adapt to specific diets and microhabitats. We explored this prediction in two ways. First, we determined the average number of OTUs shared between randomly sampled pairs of pre-settlement individuals compared to post-settlement individuals. Contrary to our expectation, the average count of shared OTUs did not vary with life stage (22.8 vs 23.2 for pre- and post-settlement, respectively). Second, we compared the richness and composition of the core microbiome between life stages, following normalization to a uniform sample size (# of host individuals) per stage. Consistent with a recent analysis of microbiomes from 15 coastal fish species (Givens et al. 2015), no single microbial OTU was present in all individuals of either group (pre or post) in our study. Contrary to our prediction, the richness of the core microbiome was similar for both pre and post-settlement stages; for example, 17 OTUs were shared across 70% of pre-settlement individuals, compared to 15 for post-settlement (Figure 3.4). Furthermore, many of the same OTUs were present in the core microbiomes identified at each stage. For example, of the OTUs occurring in 70% of total microbiome samples (pre- and post combined), 12 (of 16 total) were present in both the pre- and post-settlement core sets (70% threshold) when evaluated separately (Figure 3.4), indicating conservation of a subset of key microbial members across settlement. Together, these results suggest processes affecting assembly and diversity of the core microbiome, when defined at the OTU level, are relatively constant from pre- to post-settlement.

Many of these core OTUs were absent from or at much lower abundances in seawater samples (with the exception of 3 OTUs), indicating selective accumulation in

the fish gut (Figure 3.4). Notably a single OTU with high sequence similarity to *Vibrio harveyi* (Vibrionaceae, Genbank accession KX380754.1) comprised an average of 18-20% of the total sequence reads in both pre- and post-settlement fish, but only 2% of the seawater community (Figure 3.4). *Vibrios* are often described as opportunistic pathogens (Karunasagar et al. 1994; Austin & Zhang 2006). For example, metagenome sequences related to *Vibrio* species, including *V. harveyi*, from the guts of farmed adult turbot were enriched in genes encoding potential virulence functions (Xing et al. 2013). However, the wide distribution of *V. harveyi* in healthy marine hosts, observed here and in other studies of diverse invertebrates and fish (Onarheim et al. 1994; Guerrero-Ferreira et al. 2013;), suggests that the primary niche of this bacterium may be commensal or even mutualistic with the host, potentially with a role in protein degradation and digestion (Ray et al. 2012). Interestingly, the Vibrionaceae family as a whole (multiple OTUs) was over twice as abundant in post-settlement fish (Table 3.1) and therefore may be more typical of mature gut microbiome communities.

The proportional abundance of other core microbiome taxa varied significantly between life stages. The 16 taxa detected in 70% of all microbiome samples, a core set dominated by OTUs belonging to the Gammaproteobacteria (Figure 3.4), constituted ~60% of the total gut microbiome sequence reads in pre-settlement fish but only ~25% in settled fish, indicating that these OTUs may be important early in life. Notably, OTUs closely related to *Shewanella putrafaciens* (Genbank accession KP967510.1) and to an uncultured bacterium of the family Endozoicomonaceae (Genbank accession LN626318.1) constituted 8 and 18% of the pre-settlement microbiome, respectively, but less than 1% of the post-settlement community (Figure 3.4). To our knowledge, this is

the first report of Endozoicomonaceae occupying a large proportion of the fish gut microbiome. Bacteria of this family, within the widely distributed Oceanospiralles group, exhibit diverse heterotrophic metabolisms (Neave et al. 2014) and are commonly identified as symbionts in marine invertebrates (Bayer et al. 2013a; Bayer et al. 2013b; Forget & Juniper 2013; Nishijima et al. 2013; Beinart et al. 2014; Hyun et al. 2014), suggesting invertebrate prey as a potential vector of transport into the fish microbiome. The functional properties and potential for long-term residency of these bacteria in reef fishes remain to be verified. In contrast, Gammaproteobacteria of the genus *Shewanella*,have been detected widely in fish. Indeed, *Shewanella putrefaciens*, the species most closely related (97%) to the core OTU in this study, has been used as a probiotic to increase growth and pathogen resistance in juvenile sole (Lobo et al. 2014).

Other members of the core OTU set spanned diverse bacterial divisions, but occurred at lower abundances (Figure 3.4). These included members of genera containing known fish pathogens, such as *Photobacterium* (Chabrillon et al. 2005), *Pseudoalteromonas* (Pujalte et al. 2004), *Pseudomonas* (Wakabayashi et al. 1996), and *Halomonas* (Austin 2005). However, these groups also contain potentially beneficial members, as certain *Pseudoalteromonas* and *Alteromonas* strains can inhibit bacterial pathogens in fish (Gatesoupe 1999) and *Photobacterium* species are known bioluminescent symbionts in diverse fish hosts (Ruby & Nealson 1976; Ast et al. 2007). Core OTUs also included a Chromatiales-affiliated taxon with 100% similarity to the uncultured gill symbiont of *Ifremeria nautilei* (Beinart et al. 2013, Genbank accession KF780855.1), a member of the *Kordia* genus resembling a heterotrophic isolate from the gut of marine polychaetes (Choi et al. 2011, Genbank accession NR_117471.1), and

Salinivibrio costicola, a halotolerant, facultatively anaerobic bacterium first isolated from a hypersaline pond (Huang et al. 2000). Three OTUS matching marine Rhodobacteraceae, a broadly distributed aquatic group that has also been detected in reef surgeonfish (Miyake et al. 2015), were also prevalent across samples. However, two of these OTUs were more abundant in water samples, suggesting these taxa may be transient members of the gut community. Together, these data identify a relatively species-poor (e.g., 16 OTUs at 70%) but abundant (35-65% of total 16S sequences) core microbiome across the damselfish and cardinalfish sample set, including potentially pathogenic or beneficial bacteria, and a subset of core OTUs (12 of 16) that persist across pre- and post-settlement life stages. Persistence over development may be an important indicator of commensal or mutualistic taxa ubiquitous to reef fish.

3.5 Conclusion

These results describe a diverse gut microbiome in two abundant reef fish families and highlight the importance of life stage in structuring microbiome composition. A trend toward greater microbiome diversity in settled (older) individuals was observed in the pooled dataset, potentially explained by the hypothesis that pre-settlement fish of diverse species are more similar to each other physiologically compared to adults, occupy the same niche (pelagic zone), and acquire gut microbes from a common environmental

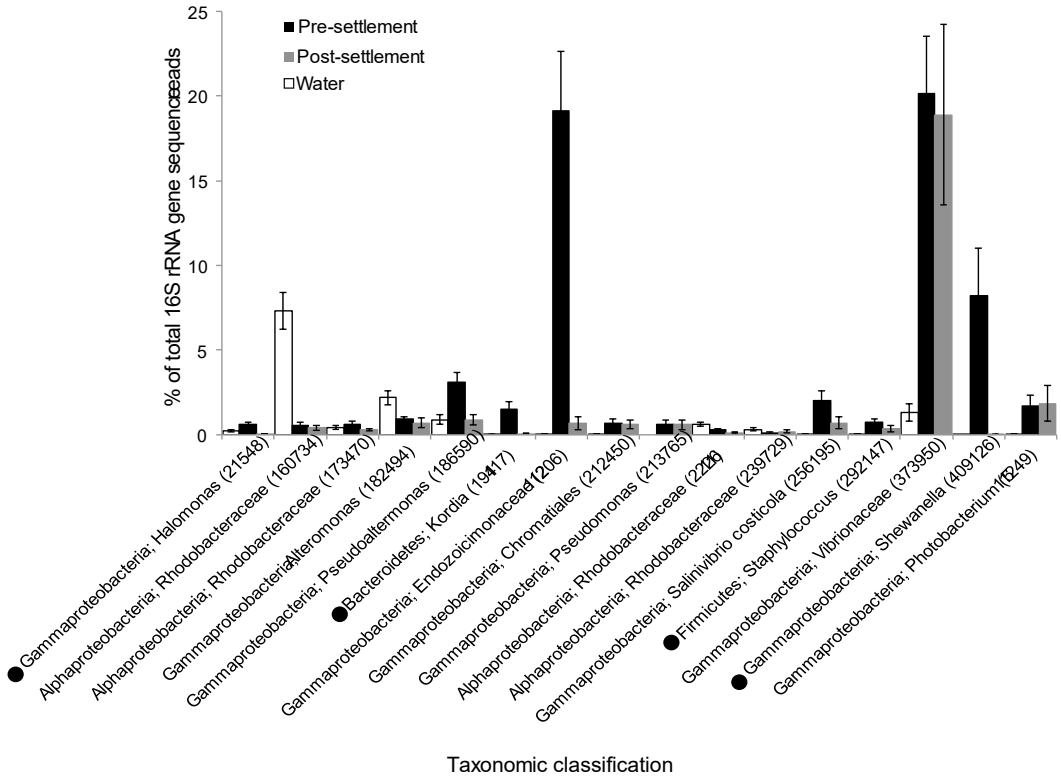

Figure 3.4. Taxonomic identity and relative abundance of OTUs detected in greater than 70% of all fish microbiome samples (pre-settlement (n=49) and post-settlement (n=24) samples) combined for analysis. □ indicates that an OTU was shared among 70% of the pre-settlement samples when this sample set was evaluated independently, but was not detected in >70% of samples in the post-settlement set when evaluated independently. The relative abundances of these OTUs in the water column samples (n=11) are included for comparison. Taxonomic classifications are at the level of genus whenever possible. Error bars are ± 1 standard error.

pool, perhaps due to similarities in diet. These factors could homogenize the pre-settlement microbiome across diverse host species, with microbiomes diversifying after fish settle on the reef and transition to adult feeding roles. However, this pattern is not uniform across host species, suggesting a need for discretion in concluding general trends in microbiome succession over development. Our data also identify core microbes common to pre- and post-settlement fish of several host species. Persistence of these

microbes over major life transitions, regardless of variable trends in bulk microbiome complexity, may implicate these microbes as particularly important to host health and physiology. Differences in microbiome assembly between pre and post-settlement fish are likely driven by a combination of factors, including physiological changes associated with development, changes in diet and potentially feeding frequency, and varying connectivity with microbes from the external environment. More research is needed to disentangle the relative contributions of these determinants over the settlement transition.

3.6 Acknowledgements

We thank the staff of Lizard Island Research Station for logistical support.

3.7 References

Al-Harbi, A. H., & M. Naim Uddin. (2004). Seasonal variation in the intestinal bacterial flora of hybrid tilapia (Oreochromis niloticus×Oreochromis aureus) cultured in earthen ponds in Saudi Arabia. *Aquaculture,* 229, pp. 37-44.

Allen, G. R. (2014). Review of Indo-Pacific coral reef fish systematics: 1980 to 2014. *Ichthyological Research*, 62(1), pp. 2-8.

Angert ER, Clements KD, & Pace, NR *(*1993). The Largest Bacterium. *Nature,* 362, *pp.* 239–241.

Austin, B. (2005). Bacterial pathogens of marine fish. In *Oceans and health: pathogens in the marine environment*, pp. 391-413, Springer US.

Austin, B., & Zhang, X. H. (2006). Vibrio harveyi: a significant pathogen of marine vertebrates and invertebrates. *Letters in applied microbiology*, *43*(2), pp. 119-124.

Ast, J. C., Cleenwerck, I., Engelbeen, K., Urbanczyk, H., Thompson, F. L., De Vos, P., & Dunlap, P. V. (2007). *Photobacterium kishitanii* sp. nov., a luminous marine bacterium symbiotic with deep-sea fishes. *International journal of systematic and evolutionary microbiology*, *57*(9), pp. 2073-2078.

Backhed, F., Ley, R., Sonnenburg, J., Peterson, D., & Gordon, J. (2005). Host-Bacterial Mutualism in the Human Intestine. *Science*, 307, pp. 1915-1920.

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

Bano, N., A. Derae Smith, W. Bennett, L. Vasquez, & J. T. Hollibaugh. (2007). Dominance of Mycoplasma in the guts of the Long-Jawed Mudsucker, *Gillichthys mirabilis*, from five California salt marshes. *Environmental Microbiology* 9, pp. 2636-2641.

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

Bates J. M., Akerlund, J., Mittge, E., & Guillemin, K. (2007). Intestinal alkaline phosphatase detoxifies lipopolysaccharide and prevents inflammation in zebrafish in response to the gut microbiota. *Cell Host Microbe* 2, pp. 371–382.

Bayer, T., Neave, M. J., Alsheikh-Hussain, A., Aranda, M., Yum, L. K., Mincer, T., ... & Voolstra, C. R. (2013a). The microbiome of the Red Sea coral Stylophora pistillata is dominated by tissue-associated Endozoicomonas bacteria. *Applied and environmental microbiology*, *79*(15), pp. 4759-4762.

Bayer T, Arif C, Ferrier-Pagès C, Zoccola D, Aranda M, Voolstra CR. (2013b). Bacteria of the genus Endozoicomonas dominate the microbiome of the Mediterranean gorgonian coral *Eunicella cavolini*. *Mar. Ecol. Prog. Ser.* 479, pp. 75–84.

Beinart, R. A., Nyholm, S. V., Dubilier, N., & Girguis, P. R. (2014). Intracellular Oceanospirillales inhabit the gills of the hydrothermal vent snail Alviniconcha

with chemosynthetic, γ-Proteobacterial symbionts. *Environmental microbiology reports*, 6(6), pp. 656-664.

Bellwood DR, Hughes TP, Folke C, & Nyström M. (2004). Confronting the coral reef crisis. *Nature*. 429, pp. 827- 833.

Beukers-Stewart, B. D., & Jones, G. P. (2004). The influence of prey abundance on the feeding ecology of two piscivorous species of coral reef fish. *Journal of Experimental Marine Biology and Ecology*, *299*(2), pp. 155-184.

Bolnick DI, Snowberg LK, Hirsch PE, Lauber CL, Knight R, Caporaso JG, & Svanbäck R. (2014a). Individuals' diet diversity influences gut microbial diversity in two freshwater fish (threespine stickleback and Eurasian perch. *Ecol Lett*., 17(8), pp. 979-87.

Bolnick, D. I., Snowberg, L. K., Hirsch, P. E., Lauber, C. L., Parks, B., Lusis, A. J., ... & Svanbäck, R. (2014b). Individual diet has sex-dependent effects on vertebrate gut microbiota. *Nature Communications*, *5*.

Bourne, D. G., Garren, M., Work, T. M., Rosenberg, E., Smith, G. W., & Harvell, C. D. (2009). Microbial disease and the coral holobiont. *Trends in microbiology*, *17*(12), pp. 554-562.

Brooker, R. M., Munday, P. L., & Ainsworth, T. D. (2010). Diets of coral-dwelling fishes of the genus Gobiodon with evidence of corallivory. *Journal of fish biology*, 76(10), pp. 2578-2583.

Brooker, R. M., Munday, P. L., Chivers, D. P., & Jones, G. P. (2015). You are what you eat: diet-induced chemical crypsis in a coral-feeding reef fish. *Proceedings of the Royal Society of London B: Biological Sciences*, *282*, p. 1799.

Burkepile DE, & Hay ME. (2008). Herbivore species richness and feeding complementarity affect community structure and function on a coral reef. *Proceedings of the National Academy of Sciences*, 105, p. 16201.

Burns, A.R., Zac Stephens, W., Stagaman, K., Wong, S., Rawls, J.F., Guillemin, K., & Bohannan, B.J. (2015). Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME J*, 10, pp. 655-664.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., ... & Huttley, G. A. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, *7*(5), pp. 335-336.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., ... & Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, *108*(Supplement 1), pp. 4516-4522.

Chabrillón, M., Rico, R. M., Balebona, M. C., & Moriñigo, M. A. (2005). Adhesion to sole, Solea senegalensis Kaup, mucus of microorganisms isolated from farmed

fish, and their interaction with *Photobacterium damselae* subsp. *piscicida*. *Journal of fish diseases*, *28*(4), pp. 229-237.

Cheesman S. & Guillemin K. (2007). We know you are in there: conversing with the indigenous gut microbiota. *Res Microbiol* 158, pp. 2–9.

Cheesman S. E., Neal, J. T., Mittge, E., Seredick, B. M., & Guillemin, K. (2010). Microbes and Health Sackler Coloquiium: Epethelial cell proliferation in the developing zebrafish intestine is regulated by the Wnt pathway and microbial signaling via Myd88. *Proc Natl Acad Sci*. 108, pp. 4570–4577.

Chimetto, L. A., Cleenwerck, I., Thompson, C. C., Brocchi, M., Willems, A., De Vos, P., & Thompson, F. L. (2010). Photobacterium jeanii sp. nov., isolated from corals and zoanthids. *International journal of systematic and evolutionary microbiology*, *60*(12), pp. 2843-2848.

Choi, A., Oh, H. M., Yang, S. J., & Cho, J. C. (2011). Kordia periserrulae sp. nov., isolated from a marine polychaete Periserrula leucophryna, and emended description of the genus Kordia. *International journal of systematic and evolutionary microbiology*, *61*(4), pp. 864-869.

Chung H, Pamp SJ, Hill JA, Surana NK, Edelman SM, Troy EB, Reading NC, Villablanca EJ, Wang S, Mora JR, Umesaki Y, Mathis D, Benoist C, Relman DA, & Kasper DL. (2012). Gut immune maturation depends on colonization with a host-specific microbiota. *Cell*, 149(7), pp. 1578-93.

Clements KD, Angert ER, Montgomery WL, & Choat JH. (2014). Intestinal microbiota in fishes: what's known and what's not. *Molecular Ecology*. 23, pp. 1891-1898.

Coates, D. (1980). Prey-size intake in humbug damselfish, Dascyllus aruanus (Pisces, Pomacentridae) living within social groups. *The Journal of Animal Ecology*, 49(1), pp. 335-340.

Cole, A. J., Pratchett, M. S., & Jones, G. P. (2008). Diversity and functional importance of coral-feeding fishes on tropical coral reefs. *Fish and Fisheries*, *9*(3), pp. 286-307.

Cole, A. J., Lawton, R. J., Wilson, S. K., & Pratchett, M. S. (2012). Consumption of tabular acroporid corals by reef fishes: a comparison with plant–herbivore interactions. *Functional Ecology*, *26*(2), pp. 307-316.

Cooper W. J., & Westneat M. W. (2009). Form and function of damselfish skulls: rapid and repeated evolution into a limited number of trophic niches. *BMC Evol. Biol*. 9, p. 24.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., ... & Andersen, G. L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and environmental microbiology*, *72*(7), pp. 5069-5072.

DiGiulio, D. B., Romero, R., Amogan, H. P., Kusanovic, J. P., Bik, E. M., Gotsch, F., ... & Relman, D. A. (2008). Microbial prevalence, diversity and abundance in

amniotic fluid during preterm labor: a molecular and culture-based investigation. *PloS one*, *3*(8), e3056.

DiGiulio DB. (2012). Diversity of microbes in amniotic fluid. *Semin Fetal Neonatal Med* pp. 17, pp. 2–11.

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

Dufrene, M., & Legendre, P. (1997). Species assemblages and indicator species: the need for a flexible assymetrical approach. *Ecological Monographs*, 67, pp. 345-366.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, pp. 2460-2461.

Emery, A. R. (1973). Comparative ecology and functional osteology of fourteen species of damselfish (Pisces: Pomacentridae) at Alligator Reef, Florida Keys. *Bulletin of Marine Science*, *23*(3), pp. 649-770.

Emery, A. R. & Thresher, R. E. (Eds) (1980). Biology of the damselfishes. *Bulletin of Marine Science* 30, pp. 145–328.

Forget NL, & Kim Juniper S. (2013). Free-living bacterial communities associated with tubeworm (*Ridgeia piscesae*) aggregations in contrasting diffuse flow hydrothermal vent habitats at the Main Endeavour Field, Juan de Fuca Ridge. *MicrobiologyOpen* 2, pp. 259–275.

Fortunato, C. S., Eiler, A., Herfort, L., Needoba, J. A., Peterson, T. D., & Crump, B. C. (2013). Determining indicator taxa across spatial and seasonal gradients in the Columbia River coastal margin. *The ISME journal*, *7*(10), pp. 1899-1911.

Franchini, P., Fruciano, C., Frickey, T., Jones, J. C., & Meyer, A. (2014). The gut microbial community of Midas Cichlid fish in repeatedly evolved Limnetic-Benthic species pairs. *PloS one*, *9*(4), e95027.

Funkhouser, L. J., & Bordenstein, S. R. (2013). Mom knows best: the universality of maternal microbial transmission. *PLoS biology*, *11*(8), e1001631.

Ganesh, S., Parris, D.J., DeLong, E.F., & Stewart, F.J. (2014). Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *ISME J* 8, pp. 187-211.

Gatesoupe, F. J. (1999). The use of probiotics in aquaculture. *Aquaculture*, *180*(1), pp. 147-165.

Ghanbari, M., Kneifel, W., & Domig, K. J. (2015). A new view of the fish gut microbiome: advances from next-generation sequencing. *Aquaculture*, *448*, pp. 464-475.

Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, Gordon JI, Relman DA, Fraser-Liggett CM, Nelson KE (2006). Metagenomic analysis of the human distal gut microbiome. *Science* 312, pp. 1355–1359.

Givens, C. E., Ransom, B., Bano, N., & Hollibaugh, J. T. (2015). Comparison of the gut microbiomes of 12 bony fish and 3 shark species. *Marine Ecology Progress Series*, *518*, pp. 209-223.

Glad T, Kristiansen V, Nielsen K, Brusetti L, Wright AD, & Sundset M. (2010). Ecological characterisation of the colonic microbiota in Arctic and sub-Arctic seals. *Microb. Ecol*. 60, pp. 320–330.

Gourlay, R. N., WYLD, S. G., & Leach, R. H. (1977). *Mycoplasma alvi*, a new species from bovine intestinal and urogenital tracts. *International Journal of Systematic Bacteriology*, *27*(2), pp. 86-96.

Gray, M. A., Stone, R. P., McLaughlin, M. R., & Kellogg, C. A. (2011). Microbial consortia of gorgonian corals from the Aleutian islands. *FEMS microbiology ecology*, *76*(1), pp. 109-120.

Greenblum, S., Turnbaugh, P., & Borenstein, E. (2012). Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proc. Natl. Acad. Sci.* 109, pp. 594–599.

Grutter, A. S. (2002). Cleaning symbioses from the parasites' perspective. *Parasitology*, 124(07), pp. 65-81.

Guerrero-Ferreira R, Gorman C, Chavez AA, Willie S, & Nishiguchi MK, (2013). Characterization of the bacterial diversity in Indo-West Pacific loliginid and sepiolid squid light organs. *Microb Ecol*., 65(1), pp. 214-26.

Hamdi, C., Balloi, A., Essanaa, J., Crotti, E., Gonella, E., Raddadi, N., ... & Cherif, A. (2011). Gut microbiome dysbiosis and honeybee health. *Journal of applied entomology*, *135*(7), pp. 524-533.

Hehemann, J. H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., & Michel, G. (2010). Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature*, 464(7290), pp. 908-912.

Hehemann, J. H., Kelly, A. G., Pudlo, N. A., Martens, E. C., & Boraston, A. B. (2012). Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. *Proceedings of the National Academy of Sciences*, *109*(48), pp. 19786-19791.

Heijtz, R., Wang, S., Anuar, F., Qian, Y., Björkholm, B., Samuelsson, A., Hibberd, M., Forssberg, H., & Pettersson, S. (2011). Normal gut microbiota modulates brain development and behavior. *Proc. Natl. Acad. Sci.* 108, pp. 3047–3052.

Hennersdorf, P., Kleinertz, S., Theisen, S., Abdul-Aziz, M. A., Mrotzek, G., Palm, H. W., & Saluz, H. P. (2016). Microbial diversity and parasitic load in tropical fish of different environmental conditions. *PloS one*, *11*(3), e0151594.

Hildebrandt, M. A., Hoffmann, C., Sherrill–Mix, S. A., Keilbaugh, S. A., Hamady, M., Chen, Y. Y., ... & Wu, G. D. (2009). High-fat diet determines the composition of the murine gut microbiome independently of obesity. *Gastroenterology*, 137(5), pp. 1716-1724.

Hinestrosa, F., Madeira, R. G., & Bourbeau, P. P. (2007). Severe gastroenteritis and hypovolemic shock caused by Grimontia (Vibrio) hollisae infection. *Journal of clinical microbiology*, *45*(10), pp. 3462-3463.

Holben, W. E., P. Williams, M. Saarinen, L. K. Sarkilahti, & J. H. A. Apajalahti. (2002). Phylogenetic analysis of intestinal microflora indicates a novel mycoplasma phylotype in farmed and wild salmon. *Microb. Ecol*. 44: pp. 175-185.

Huang, C. Y., Garcia, J. L., Patel, B. K., Cayol, J. L., Baresi, L., & Mah, R. A. (2000). Salinivibrio costicola subsp. vallismortis subsp. nov., a halotolerant facultative anaerobe from Death Valley, and emended description of Salinivibrio costicola. *International journal of systematic and evolutionary microbiology*,*50*(2), pp. 615-622.

Hyun DW, Shin NR, Kim MS, Oh SJ, Kim PS, Whon TW, & Bae JW. (2014). Endozoicomonas atrinae sp. nov., a bacterium isolated from the intestine of a comb pen shell Atrina pectinata. *Int. J. Syst. Evol. Microbiol*. 64, pp. 2312–2318.

Ingerslev, H. C., von Gersdorff Jørgensen, L., Strube, M. L., Larsen, N., Dalsgaard, I., Boye, M., & Madsen, L. (2014). The development of the gut microbiota in rainbow trout (Oncorhynchus mykiss) is affected by first feeding and diet type. *Aquaculture*, *424*, pp. 24-34.

Jiménez, E., Fernández, L., Marín, M. L., Martín, R., Odriozola, J. M., Nueno-Palop, C., ... & Rodríguez, J. M. (2005). Isolation of commensal bacteria from umbilical cord blood of healthy neonates born by cesarean section. *Current microbiology*, *51*(4), pp. 270-274.

Johnson, W. S., & Ruben, P. (1988). Cleaning behavior of Bodianus rufus, Thalassoma bifasciatum, Gobiosoma evelynae, and Periclimenes pedersoni along a depth gradient at Salt River Submarine Canyon, St. Croix. *Environmental Biology of Fishes*, 23(3), pp. 225-232.

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

Karunasagar, I., Pai, R., Malathi, G. R., & Karunasagar, I. (1994). Mass mortality of Penaeusmonodon larvae due to antibiotic-resistant Vibrio harveyi infection. *Aquaculture*, *128*(3), pp. 203-209.

Kavanagh, K. D. (2000). Larval brooding in the marine damselfish Acanthochromis polyacanthus (Pomacentridae) is correlated with highly divergent morphology, ontogeny and life-history traits. *Bulletin of Marine Science*, *66*(2), pp. 321-337.

Kimes, N. E., Johnson, W. R., Torralba, M., Nelson, K. E., Weil, E., & Morris, P. J. (2013). The Montastraea faveolata microbiome: ecological and temporal

influences on a Caribbean reef-building coral in decline. *Environmental microbiology*, *15*(7), pp. 2082-2094.

Klumpp, D. W., & Polunin, N. V. C. (1989). Partitioning among grazers of food resources within damselfish territories on a coral reef. *Journal of Experimental Marine Biology and Ecology*, 125(2), pp. 145-169.

Kokita, T., & Nakazono, A. (2001). Rapid response of an obligately corallivorous filefish Oxymonacanthus longirostris (Monacanthidae) to a mass coral bleaching event. *Coral Reefs*, 20(2), pp. 155-158.

Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., & Schloss, P.D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 79: pp. 5112- 5120.

Kubo, K., Lloyd, K. G., Biddle, J. F., Amann, R., Teske, A., & Knittel, K. (2012). Archaea of the Miscellaneous Crenarchaeotal Group are abundant, diverse and widespread in marine sediments. *The ISME journal*, 6(10), pp. 1949-1965.

Lee WJ, & Hase K. (2014). Gut microbiota-generated metabolites in animal health and disease. *Nat Chem Biol*., 10(6), pp. 416-24.

Leis JM (1991) The pelagic stage of reef fishes: the larval biology of coral reef fishes. In: Sale PF (ed) *The ecology of fishes on coral reefs*. Academic Press, San Diego, Calif. pp. 183–230.

Leis, J. M., Sweatman, H. P., & Reader, S. E. (1996). What the pelagic stages of coral reef fishes are doing out in blue water: daytime field observations of larval behavioural capabilities. *Marine and Freshwater Research*, *47*(2), pp. 401-411.

Ley, R., Peterson, D., & Gordon, J. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 124, pp. 837–848.

Llewellyn MS, Boutin S, Hoseinifar SH, & Derome N. (2014). Teleost microbiomes: the state of the art in their characterization, manipulation and importance in aquaculture and fisheries. *Front Microbiol*. 5, p. 207.

Llewellyn, M. S., McGinnity, P., Dionne, M., Letourneau, J., Thonier, F., Carvalho, G. R., ... & Derome, N. (2015). The biogeography of the atlantic salmon (Salmo salar) gut microbiome. *The ISME journal*.

Lobo, C., Moreno-Ventas, X., Tapia-Paniagua, S., Rodríguez, C., Moriñigo, M. A., & de La Banda, I. G. (2014). Dietary probiotic supplementation (Shewanella putrefaciens Pdp11) modulates gut microbiota and promotes growth and condition in Senegalese sole larviculture. *Fish physiology and biochemistry*, *40*(1), pp. 295-309.

Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K., & Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature*, *489*(7415), pp. 220-230.

Luczkovich, J. J., & Stellwag, E. J. (1993). Isolation of cellulolytic microbes from the intestinal tract of the pinfish, Lagodon rhomboides: size-related changes in diet and microbial abundance. *Marine Biology*, 116(3), pp. 381-388.

Magarinos, B., Romalde, J.L., Noya, M., Barja, J.L., & Toranzo, A.E., (1996). Adherence and invasive capacities of the fish pathogen Pasteurella piscicida. *FEMS Microbiol. Lett*. 138, pp. 29–34.

Magoč, T., & Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27: pp. 2957-2963.

Macián, M. C., Garay, E., Grimont, P. A. D., & Pujalte, M. J. (2004). *Vibrio ponticus* sp. nov., a neighbour of *V. fluvialis-V. furnissii* clade, isolated from gilthead sea bream, mussels and seawater. *Systematic and applied microbiology*, *27*(5), pp. 535-540.

Marnane, M. J., & Bellwood, D. R. (2002). Diet and nocturnal foraging in cardinalfishes (Apogonidae) at one tree reef, great barrier reef, Australia. *Marine Ecology progress series*, 231, pp. 261-268.

Martens, E. C., Kelly, A. G., Tauzin, A. S., & Brumer, H. (2014). The devil lies in the details: how variations in polysaccharide fine-structure impact the physiology and evolution of gut microbes. *Journal of molecular biology*, 426(23), pp. 3851-3865.

Martin-Antonio, B., Manchado, M., Infante, C., Zerolo, R., Labella, A., Alonso, C., & Borrego, J. J. (2007). Intestinal microbiota variation in Senegalese sole (Solea senegalensis) under different feeding regimes. *Aquaculture research*, 38: pp. 1213-1222.

McCauley, D. J., Micheli, F., Young, H. S., Tittensor, D. P., Brumbaugh, D. R., Madin, E. M., ... & Worm, B. (2010). Acute effects of removing large fish from a near-pristine coral reef. *Marine biology*, 157, pp. 2739.

Miyake, S., Ngugi, D. K. & Stingl, U. (2015), Diet strongly influences the gut microbiota of surgeonfishes. *Molecular Ecology*, 24, pp. 656–672.

Moeller AH, Li Y, Mpoudi Ngole E, Ahuka-Mundeke S, Lonsdorf EV, Pusey AE, Peeters M, Hahn BH, & Ochman H. (2014). Rapid changes in the gut microbiome during human evolution. *Proc Natl Acad Sci*, 111(46), pp. 16431-5.

Mouchet, M. A., Bouvier, C., Bouvier, T., Troussellier, M., Escalas, A., & Mouillot, D. (2012). Genetic difference but functional similarity among fish gut bacterial communities through molecular and biochemical fingerprints. *FEMS microbiology ecology*, *79*(3), pp. 568-580.

Mumby P. J., Dahlgren, C. P., Harborne, A. R., Kappel, C. V., Micheli, F., Brumbaugh, D. R., ... & Gill, A. B., (2006). Fishing, trophic cascades, and the process of grazing on coral reefs. *Science* 311, p. 98.

Nayak SK (2010). Role of gastrointestinal microbiota in fish. *Aquaculture Research*, 41, pp. 1553–1573.

Neave, M. J., Michell, C. T., Apprill, A., & Voolstra, C. R. (2014). Whole-genome sequences of three symbiotic Endozoicomonas strains. *Genome announcements*, *2*(4), e00802-14.

Nishijima M, Adachi K, Katsuta A, Shizuri Y, & Yamasato K. (2013). Endozoicomonas numazuensis sp. nov., a gammaproteobacterium isolated from marine sponges, and emended description of the genus Endozoicomonas Kurahashi and Yokota 2007. *Int. J. Syst. Evol. Microbiol*. 63, pp. 709–714.

Onarheim, A. M., Wiik, R., Burghardt, J., & Stackebrandt, E. (1994). Characterization and identification of two Vibrio species indigenous to the intestine of fish in cold sea water; description of Vibrio iliopiscarius sp. nov. *Systematic and applied microbiology*, *17*(3), pp. 370-379.

Papazisi, L., Gorton, T. S., Kutish, G., Markham, P. F., Browning, G. F., Swartzell, S., ... & Geary, S. J. (2003). The complete genome sequence of the avian pathogen Mycoplasma gallisepticum strain Rlow. *Microbiology*, *149*(9), pp. 2307-2316.

Pascual, J., Macián, M. C., Arahal, D. R., Garay, E., & Pujalte, M. J. (2009). Description of *Enterovibrio nigricans* sp. nov., reclassification of *Vibrio calviensis* as *Enterovibrio calviensis* comb. nov. and emended description of the genus Enterovibrio Thompson et al. 2002. *International journal of systematic and evolutionary microbiology*, *59*(4), pp. 698-704.

Patton, W.K. Community structure among the animals inhabiting the coral Pocillopora damicornis at Heron Island, Australia (1974). In *Symbiosis in the Sea*; Vernberg, W., Ed.; University of South Carolina Press: Columbia, SC, USA, pp. 219–243.

Pollak, P.E. & W.L. Montgomery. (1994). Giant bacterium (Epulopiscium fishelsoni) influences digestive enzyme activity of an herbivorous surgeonfish (Acanthurus nigrofuscus). *Comp. Biochem. Physiol*, 108A. pp. 657-662.

Pratchett, M. S. (2007). Dietary selection by coral-feeding butterflyfishes (Chaetodontidae) on the Great Barrier Reef, Australia. *Raffles Bull Zool*, 14, pp. 171-176.

Pujalte Domarco, M. J., Sitjà Bobadilla, A., Álvarez Pellitero, P., & Garay Auban, E. (2004). Carriage of potentially fish-pathogenic bacteria in Sparus aurata cultured in Mediterranean fish farms. *Diseases of Aquatic Organisms*, 54, pp. 119-126.

Randall, J. E. (1967). Food habits of reef fishes of the West Indies. Institute of Marine Sciences, University of Miami.

Rawls J., Samuel, B. S., & Gordon, J. I., (2004). Gnotobiotic zebrafish reveal evolutionarily conserved responses to the gut microbiota. *Proc Natl Acad Sci*, 101, pp. 4596–4601.

Ray, A. K., Ghosh, K., & Ringø, E. (2012). Enzyme-producing bacteria isolated from fish gut: a review. *Aquaculture Nutrition*, 18(5), pp. 465-492.

Reece, R. M., Cunha, C. B., & Rich, J. D. (2014). *Corynebacterium minutissimum* vascular graft infection: Case report and review of 281 cases of prosthetic device-

related Corynebacterium infection. *Scandinavian journal of infectious diseases*, 46(9), pp. 609-616.

Roeselers, G., Mittge, E. K., Stephens, W. Z., Parichy, D. M., Cavanaugh, C. M., Guillemin, K., & Rawls, J. F. (2011). Evidence for a core gut microbiota in the zebrafish. *The ISME journal*, *5*(10), pp. 1595-1608.

Romalde, J. L. (2002). Photobacterium damselae subsp. piscicida: an integrated view of a bacterial fish pathogen. *International Microbiology*, 5(1), pp. 3-9.

Rotjan RD, and Lewis SM (2008). Impact of coral predators on tropical reefs. *Mar Ecol Prog Ser* 367, pp. 73-91.

Round, J. & Mazmanian S. (2009). The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol*, 9, pp. 313–323.

Ruby, E. G., & Nealson, K. H. (1976). Symbiotic association of Photobacterium fischeri with the marine luminous fish Monocentris japonica: a model of symbiosis based on bacterial studies. *The Biological Bulletin*, 151(3), pp. 574-586.

Sanchez, L. M., Wong, W. R., Riener, R. M., Schulze, C. J., & Linington, R. G. (2012). Examining the fish microbiome: vertebrate-derived bacteria as an environmental niche for the discovery of unique marine natural products. *PloS one*, 7(5), e35398.

Schmale, M. C., Hensley, G. T., & Udey, L. R. (1986). Neurofibromatosis in the Bicolor Damselfish (Pomacentrus partitus) as a Model of von Recklinghausen Neurofibromatosisa. *Annals of the New York Academy of Sciences*, *486*(1), pp. 386-402.

Schmidt, V.T., Smith, K.F., Melvin, D.W., & Amaral-Zettler, L.A. (2015). Community assembly of a euryhaline fish microbiome during salinity acclimation. *Molecular Ecology*, 24, pp 2537-2550.

Shade A, & Handelsman J. (2012). Beyond the Venn diagram: the hunt for a core microbiome. *Environ Microbiol*., 14(1), pp. 4-12.

Smriga, S., Sandin, S., & Azam, F. (2010). Abundance, diversity, and activity of microbial assemblages associated with coral reef fish guts and feces. *FEMS Microbiol Ecol*, 73, pp. 31-42.

Spanggaard, B., I. Huber, J. Nielsen, T. Nielsen, K. F. Appel, & L. Gram. (2000). The microflora of rainbow trout intestine: a comparison of traditional and molecular identification. *Aquaculture*, 182, pp. 1-15.

Sudakaran, S., Salem, H., Kost, C., & Kaltenpoth, M. (2012). Geographical and ecological stability of the symbiotic mid-gut microbiota in European firebugs, Pyrrhocoris apterus (Hemiptera, Pyrrhocoridae). *Molecular Ecology*, 21(24), pp. 6134-6151.

Sullam, K.E., Essinger, S. D., Lozupone, C. A., O'CONNOR, M. P., Rosen, G. L., Knight, R. O. B., ... & Russell, J. A. (2012). Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. *Molecular Ecology*, 21, pp. 3363-3378.

Sullam, K. E., Rubin, B. E., Dalton, C. M., Kilham, S. S., Flecker, A. S., & Russell, J. A. (2015). Divergence across diet, time and populations rules out parallel evolution in the gut microbiomes of Trinidadian guppies. *The ISME journal*, 9(7), pp. 1508-1522.

Theis KR, Venkataraman, A., Dycus, J. A., Koonter, K. D., Schmitt-Matzen, E. N., Wagner, A. P., ... & Schmidt, T. M. (2013). Symbiotic bacteria appear to mediate hyena social odors. *Proc Natl Acad Sci.*, 110, 49, pp. 19832–19837.

Thompson, F. L., Thompson, C. C., Naser, S., Hoste, B., Vandemeulebroecke, K., Munn, C., ... & Swings, J. (2005). Photobacterium rosenbergii sp. nov. and Enterovibrio coralii sp. nov., vibrios associated with coral bleaching. *International journal of systematic and evolutionary microbiology*, *55*(2), pp. 913-917.

Turnbaugh, P.J., Ley, R. E., Mahowald, M. A., Magrini, V., Mardis, E. R., & Gordon, J. I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest, *Nature*, 444, pp. 1027-1031.

Turnbaugh PJ, Ridaura, V. K., Faith, J. J., Rey, F. E., Knight, R., & Gordon, J. I. (2009a). The effect of diet on the human gut microbiome: A metagenomic analysis in humanized gnotobiotic mice. *Sci Transl Med* 1,6, 6ra14.

Turnbaugh, P. J., & Gordon, J. I. (2009b). The core gut microbiome, energy balance and obesity. *The Journal of physiology*, 587(17), pp. 4153-4158.

Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, & Gordon JI. (2009c). A core gut microbiome in obese and lean twins. *Nature*, 457(7228), pp. 480-484.

Van Der Maarel, M. J., Artz, R. R., Haanstra, R., & Forney, L. J. (1998). Association of marine archaea with the digestive tracts of two marine fish species. *Applied and environmental microbiology*, *64*(8), pp. 2894-2898.

Vasconcelos, A. T. R., Ferreira, H. B., Bizarro, C. V., Bonatto, S. L., Carvalho, M. O., Pinto, P. M., ... & Martinkovics, C. T. (2005). Swine and poultry pathogens: the complete genome sequences of two strains of Mycoplasma hyopneumoniae and a strain of Mycoplasma synoviae. *Journal of bacteriology*, *187*(16), pp. 5568-5577.

Vetriani C, Jannasch HW, MacGregor BJ, Stahl DA, & Reysenbach AL. (1999). Population structure and phylogenetic characterization of marine benthic Archaea in deep-sea sediments. *Appl Environ Microbiol*., 10, pp. 4375-84.

Wada-Katsumata A, Zurek L, Nalyanya G, Roelofs WL, Zhang A, & Schal C. (2015). Gut bacteria mediate aggregation in the German cockroach. *Proceedings of the National Academy of Sciences of the United States of America*, 112(51), pp. 15678-15683.

Wakabayashi H., Sawada K., Ninomiya K., & Nishimori E. (1996). Bacterial hemorrhagic ascites of ayu caused by Pseudomonas sp. *Fish Pathol*. 31, pp. 239–240.

Ward, N. L., B. Steven, K. Penn, B. A. Methé, & W. H. Detrich. (2009). Characterization of the intestinal microbiota of two Antarctic notothenioid fish species. *Extremophiles*, 13, pp. 679-685.

Wellington, G. M., & Victor, B. C. (1989). Planktonic larval duration of one hundred species of Pacific and Atlantic damselfishes (Pomacentridae). *Marine Biology*, *101*(4), pp. 557-567.

Whitman WB, Coleman DC, & Wiebe WJ (1998). Prokaryotes: the unseen majority. *Proc Natl Acad Sci* 95, pp. 6578–6583.

Wilson S., & Bellwood D. 1997. Cryptic dietary components of territorial damselfishes (Pomacentridae, Labroidei). *Mar. Ecol. Prog. Ser*. 153, pp. 299–310.

Wilson, S. K., Burgess, S. C., Cheal, A. J., Emslie, M., Fisher, R., Miller, I., ... & Sweatman, H. (2008). Habitat utilization by coral reef fish: implications for specialists vs. generalists in a changing environment. *Journal of Animal Ecology*, *77*(2), pp. 220-228.

Wong, S., & Rawls, J. F. (2012). Intestinal microbiota composition in fishes is influenced by host ecology and environment. *Molecular ecology*, *21*(13), pp. 3100-3102.

Xing, M., Hou, Z., Yuan, J., Liu, Y., Qu, Y., & Liu, B. (2013). Taxonomic and functional metagenomic profiling of gastrointestinal tract microbiome of the farmed adult turbot (Scophthalmus maximus). *FEMS Microbiology Ecology*, 86(3), pp. 432-443.

Ye L, Amberg J, Chapman D, Gaikowski M, & Liu WT. (2014). Fish gut microbiota analysis differentiates physiology and behavior of invasive Asian carp and indigenous American fish. *ISME J*., 8, pp. 541–551.

Yoshitomo, K., Xian-Ying, M., & Takema, F. (2005). Gut symbiotic bacteria of the genus Burkholderia in the broad-headed bugs Riptortus clavatus and Leptocorisa chinensis (Heteroptera: Alydidae). Applied and Environmental Microbiology, 71, pp. 4035-4043.

Zheng X, Zhao, A., Xie, G., Chi, Y., Zhao, L., Li, H., ... & Jia, W. (2013). Melamine-induced renal toxicity is mediated by the gut microbiota. *Sci Transl Med*., 5, 172, 172ra22.

CHAPTER 4.

FEEDING RAPIDLY ALTERS MICROBIOME COMPOSITION AND GENE TRANSCRIPTION IN THE CLOWNFISH GUT

Disclaimer: This chapter is currently in review at the journal of Applied and Environmental Microbiology

4.1 Abstract

Diet is a major determinant of intestinal microbiome composition. While studies have evaluated microbiome responses to diet variation, less is understood of how the act of feeding influences the microbiome, independent of diet type. Here, we use the clownfish *Premnas biaculeatus*, a species reared commonly in ornamental marine aquaculture, to test how the diversity, predicted gene content, and gene transcription of the microbiome vary over a two-day diurnal period with a single daily feeding event. This study used fish fed four times daily, once daily, or every three days prior to the diurnal period, allowing us also to test how feeding frequency affected microbiome diversity. The amount of time between feedings had no effect on baseline diversity of the microbiome. In contrast, the act of feeding itself caused a significant short-term change in the microbiome, with microbiome diversity, predicted gene content, and gene transcription varying significantly between time points immediately before and 1.5 hours post feeding. Variation was driven by abundance shifts involving exact sequence variants (ESVs), with one ESV identified as *Photobacterium* sp. increasing from <0.5% of sequences immediately pre-feeding to 34% at 1.5 hours post-feeding. Other ESVs from a range of microbial groups also increased dramatically after feeding, with the majority also detected in the food. One ESV identified as *Clostridium perfringens* represented up to 55% of sequences but did not vary significantly over the diurnal period and was not detected in the food. Post-feeding samples were enriched in transcripts and predicted genes for social interactions, cell motility, and coping with foreign DNA, whereas time points farther from feeding were enriched in genes of diverse catabolic and biosynthetic functions. These results confirm feeding as a significant destabilizing force in clownfish

intestinal microbiomes, likely due to both input of cells attached to food and stimulation of resident microbes. Microbes such as *Photobacterium* may episodically transition from environmental reservoirs to growth in the gut, likely in association with food particles. This transition may be facilitated by functions for navigating a new environment and interacting with neighboring microbes and host cells. Other taxa, such as *Clostridium,* are comparatively stable intestinal members and less likely to be affected by passing food. Conclusions about microbiome ecology may therefore differ based on when samples were collected relative to the last feeding.

4.2 Background

The biological significance of host-associated microbiomes is widely recognized, with the intestinal (gut) microbiome in particular now known to play a vital role in host health (Chung et al. 2012; Lee & Hase 2014). Factors shaping the intestinal microbiome are complex, and include host phylogeny, diet, age, and immune status (Lozupone et al. 2012, Conlon and Bird 2014). Understanding these factors is essential, as changes in microbiome composition are linked to diverse aspects of host physiology including efficiency of nutrient acquisition, development of the intestine, immune function, and cognition and behavior (Galland 2014, Lee and Hase 2014, Gilbert et al. 2016). However, for most organisms, particularly for non-mammal systems, we lack basic knowledge of how and over what timescale intestinal microbiomes change in response to perturbations, including changes in chemical availability due to feeding.

A lack of knowledge of the dynamics and drivers of microbiome change due to feeding is due partly to the fact that few studies have sampled the intestine over the course of a feeding cycle, which requires sacrificing animals at hourly timescales.

Furthermore, interpreting short-term microbiome fluctuations is challenging, as these may also be due partly to host circadian rhythms independent of feeding (Thaiss et al. 2014,2016), or potentially to changes in feeding frequency (Voight et al. 2016, Kaczmarek et al. 2017). Zarrinpar et al. (2014), for example, compared mice allowed to feed *ad libitum* to mice fed once a day and found that both groups exhibited predictable diurnal oscillations in intestinal taxa, but differed in the number of taxa exhibiting oscillations. In addition, reduced feeding frequency during times when mice were most active was correlated with lower abundance of *Lactobacillus* bacteria, a pattern associated with protection from metabolic disease (Li et al. 2013).

The act of feeding may influence microbiome composition and function through diverse mechanisms, independent of diet type. Food intake can introduce new microbes or genes to the intestine (Heheman et al. 2010). Research on probiotic use in gnotobiotic mice shows that microbes from fermented milk products are detected in stool in a matter of days after consumption (McNulty et al. 2011). However, the extent to which the intestinal microbiome is restructured by food-attached microbes in non-model animals has been largely unstudied, although hypothesized to be potentially significant (Nayak, 2010). This is surprising, as such restructuring, particularly if transient, could bias conclusions about which microbes live as residents in a stable and potentially beneficial relationship with the host.

Feeding might also stimulate the growth of microbes already present in the intestine (residents). The growth response of residents could be associated with both changes in community structure as well as metabolic cascades linked to food breakdown. Microbes and enzymes specialized for the catabolism of complex carbohydrates might be

abundant early in digestion, creating products that can be used for energy by different microbes later in digestion (Comstock and Coyne, 2003, Belenguer et al. 2006, De Vuyst and Leroy 2011). Such cross-feeding is common in the mammalian intestine and a potential determinant of microbial richness (Sung et al. 2017, van Hoek and Merks 2017). Feeding may also promote successional patterns in microbial growth or metabolism independent of metabolite exchange. The mammalian large intestine, for example, is dominated by anaerobic microbes in the mid-lumen, but also contains aerobic or microaerobic taxa, notably along the submucosal surface closer to the oxygenated blood (Espey 2013, Aldenberg et al. 2014, Donaldson et al. 2016). Similar to what has been shown in chemostat cultures (Gerritse et al. 1990), oxygen consumption by aerobes after feeding may precede or promote anaerobic metabolism by other microbes. Indeed, the progression of digestion in insects (from foregut to the hindgut) is characterized by a linear decrease in oxygen concentration (Johnson et al. 2000, Gross et al. 2008).

The frequency of feeding might also affect the magnitude of microbiome change in response to feeding. Continuous or near-continuous feeding, by grazing animals for example, may maintain relatively constant substrate conditions in the gut as well as a steady stream of food-associated microbes, and therefore a stable assemblage of microbes with slight compositional shifts post-feeding. In contrast, intermittent feeding may promote large compositional changes associated with the transition from a relatively inactive, but stable, "fasting" microbiome to a "bloom" community after feeding. In vertebrate guts, fasting/feeding cycles have been shown to drastically alter the abundance of individual bacterial groups, with fasting associated with higher occurrence of Bacteriodetes (Crawford et al. 2009, Kohl et al. 2014). In addition to altering the

magnitude of microbiome change in response to feeding, differences in feeding regime have been shown to affect baseline microbiome composition and metabolite production (Thais et al. 2014, Li et al. 2017). However, the factors driving microbiome differences linked to feeding regime are likely complex, and potentially related to changes in host physiology (Secor and Carey 2016), as well as sampling the microbiome at varying stages in the digestive cycle.

The studies mentioned above, conducted primarily in mammalian models, highlight the need to account for short-term shifts in microbiome composition in comparative studies. However, the extent to which microbiomes of other major animal groups exhibit short-term fluctuation in response to feeding, or other diurnal cues, remains largely uncharacterized. Here, we test how the act of feeding and feeding frequency affect the intestinal microbiome of maroon clownfish (*Premnas biaculeatus*). Fishes are among the most species-rich and ecologically important vertebrate groups, with diverse roles in food webs and as targets of human recreational and commercial interest. Fisheries and aquaculture are multibillion-dollar industries, yet baseline knowledge of the diversity, function, and dynamics of fish microbiomes is absent for most commercially important fish species. *Premnas biaculeatus* in particular is widely bred for the marine ornamental fish trade, is hardy during captivity and in experiments, and represents one of the most abundant and widespread Families (Pomacentridae) in tropical seas. Using this species as a model in feeding experiments, we show large changes in the diversity and presumed metabolic function of the microbiome over a daily feeding cycle but minimal prolonged effect of feeding frequency on baseline microbiome composition. We discuss these results in the context of taxon-specific differences in

microbial lifestyle and the potential for feeding events to influence conclusions about microbiome stability and ecology.

4.3 Results

We sampled a cohort of maroon clownfish (*Premnas biaculeatus*) at five time points per day over a two-day period with one feeding event per day (at 1100). Before sampling occurred, fish were divided among three treatment groups differing based on whether they were fed four times daily (4X, n=29), once daily (1X, n=29), or once every three days (0.33X, n=30) over the month prior to the two-day sampling period. Three fish per treatment group were sacrificed immediately pre-feeding (1100, n=17), 1.5 hours post-feeding (1230), 3 hours post-feeding (1400), 5 hours post-feeding (1600, n=17), and 9 hours post-feeding (2000, n=17) on each of the two days of the sampling period. Thus, each time point (1100, 1230, 1400, 1600, and 2000) was represented by 18 individuals (3 replicates per daily time point X 2 days of sampling X 3 treatment groups; for some time points, we obtained data from 17 rather than 18 individuals due to sample/fish loss). These samples were used for analysis of the intestinal microbiome along with samples of the water and food (see Methods for full experimental and sampling details).

4.3.1 Microbiome taxonomic composition

We recovered 6,732,587 16S rRNA gene amplicon sequences (range: 706-239,836 per sample; average: 64,120) and 2748 unique exact sequence variants (ESVs) across all samples of fish, tank water, and food (average: 288 ESVs per sample) (Table C.1). Microbiome composition in water and food samples differed significantly from that of the fish intestine (PCoA, ANOVA, $P<0.05$; Figure C.1). More than 90% of sequences in the water (across all replicates) were identified as belonging to the bacterial Families

Flavobacteriaceae (Bacteroidetes), Methylophilaceae (Alphaproteobacteria), or Rhodobacteraceae (Alphaproteobacteria). Food microbiomes were dominated (~70%) by sequences of the Phormidiaceae and Streptophyta (Cyanobacteria), Pseudoalteromonadaceae (Gammaproteobacteria), and Rickettsiales (Alphaproteobacteria) (Figures C.1-C.2). In contrast, fish microbiomes were composed primarily (~70%) of sequences of the Families Clostridiaceae (Firmicutes), Mycobacteriaceae (Actinobacteria), and Vibrionaceae (Gammaproteobacteria; Figure C.1). Many groups, notably Vibrionaceae, were detected in both the intestine and food samples (see below). Alpha diversity (Shannon index) did not differ significantly among water, food, and fish samples (fish vs. water vs. food, all samples combined).

Fish microbiome composition and alpha diversity during the two-day diel sampling period did not differ among samples grouped based on feeding treatment (4X, 1X, or 0.33X daily feedings in the month prior to sampling; Figure C.3). Furthermore, no ESVs were detected as differentially abundant among feeding groups. Similarly, no differences were observed when analyzing only samples from the first time point (pre-feeding, 1100) on day 1 of the diel sampling, prior to the synchronization of feeding schedules for the diel sampling, and presumably the point at which the intestinal communities of different groups would be most affected by prior feeding regime.

In contrast, microbiomes varied significantly among samples grouped based on time of day during the two-day diel sampling period, irrespective of feeding frequency treatment in the prior 30 days (Figure 4.1A, Figures C.4-C.5). Microbiomes generally partitioned into two clusters, one including nearly all samples from the two time points immediately post-feeding (1230, 1400; “fed” in Figure 4.1A), and another including the

two time points most separated from the feeding event (1100, 2000; “unfed”). Microbiomes at the intermediate time point (1600, 5 hours post-feed) were more variable, with replicates falling within both the fed and unfed clusters. Consistent with these clustering patterns, random forest analysis based on 698 ESVs (best model) showed that microbiomes from the 1100 and 1230 time points could be correctly classified to those time points with 100% accuracy, whereas ESV composition was less predictive of later time points post-feeding (Figure 4.2).

Average alpha diversity (Shannon index) also varied significantly based on sampling time, being lowest immediately pre-feeding at 1100, increasing by over 2-fold by 1.5 hours post-feeding (1230), and then decreasing steadily thereafter as intestinal content cleared (Figure 4.3). Shannon diversity at 1230 was significantly higher compared to every other time point, excluding 1600; diversity at 2000 was significantly lower compared to every other time point, excluding 1100 (Kruskal Wallis pairwise, corrected p-value <0.05).

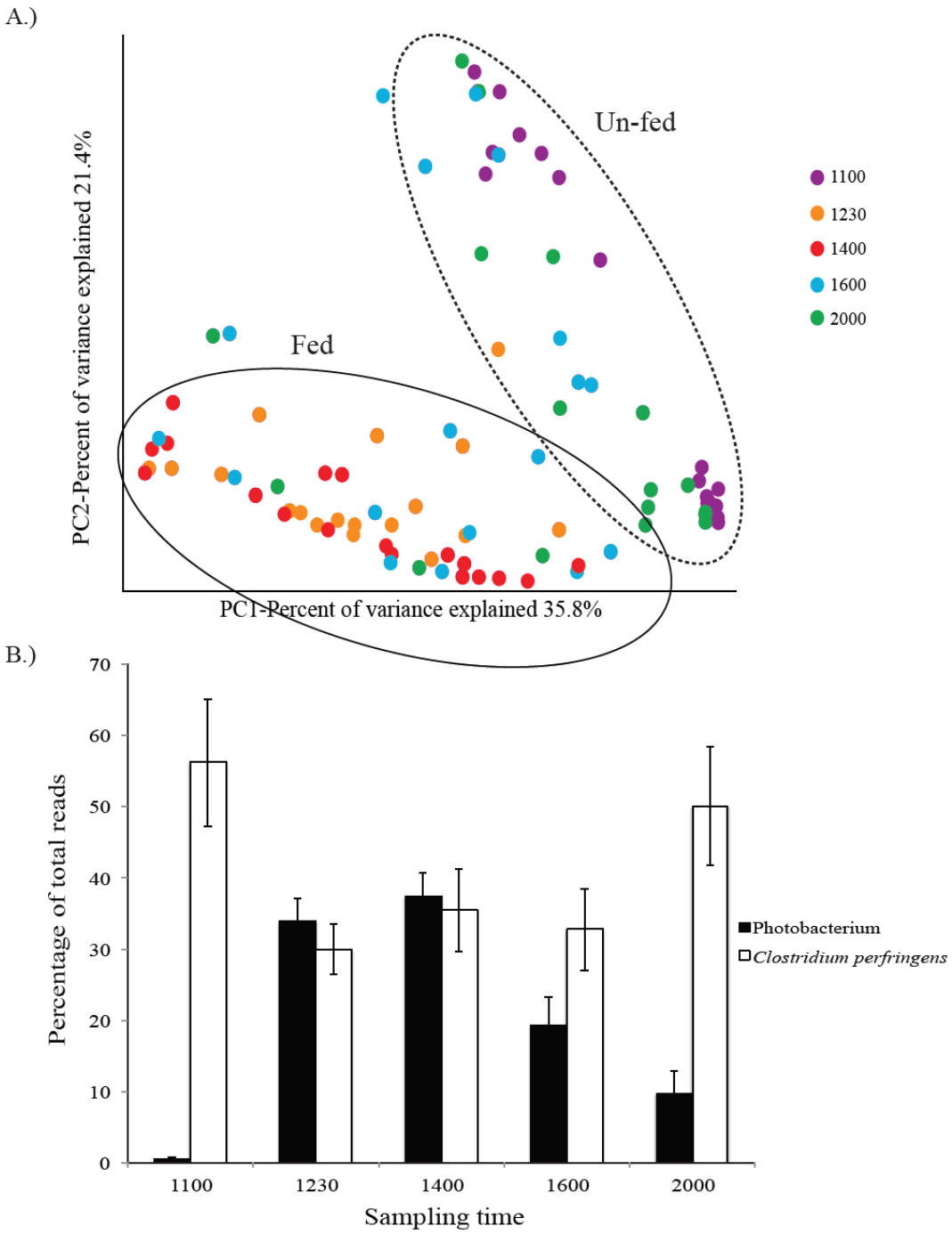

Figure 4.1. Intestinal microbiome composition varies over a feeding cycle A.) Principal coordinate analysis showing significant microbiome clustering by time of sampling (ANOVA, $P<0.05$). Distance was based on Bray-Curtis similarities with all samples rarefied to 20,562 sequences. Samples within the dashed circle were primarily from the unfed time points while samples within the solid circle were from the fed time points. B.) Relative abundance of the two most common ESVs (exact sequence variants) varies over a feeding cycle (n=17 at 1100 and 2000, n=18 at 1230, 1400, and 1600). ANCOM analysis identified the *Photobacterium* ESV, but not the *Clostridium* ESV, as varying significantly over the sampling period ($P<0.05$).

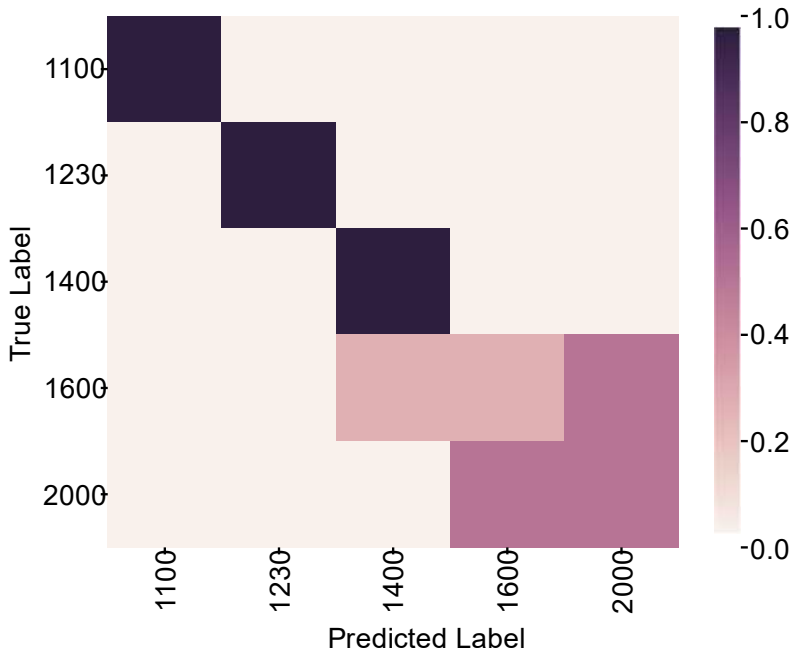

Figure 4.2. Samples can be accurately classified to sampling time based on community composition. Random forest model showing the frequency at which microbiome samples from a given time point are correctly assigned based on taxonomic composition. This model included 698 unique sequence features (best model) and had an overall accuracy of 72% (n=17 fish at 1100 and 2000 and18 fish at 1230, 1400, and 1600).

Diverse microbial groups fluctuated in abundance over the two-day sampling (Figure C.4, Table C.2). The most abundant ESV in the dataset was classified as *Clostridium perfringens* (Firmicutes). This ESV comprised 50-56% of all amplicons at unfed time points (1100, 2000) and 29-35% at the intermediate, fed time points (Figure 4.1B); however, this variation was not statistically significant (P>0.05) based on analysis of composition of microbiomes (ANCOM, Mandal et al. 2015). ANCOM analysis, however, identified 57 ESVs whose abundance varied significantly based on time (Table C.2). Almost all of these (55 of 57) increased markedly in representation from pre-feeding (1100) to 1.5 hours post-feeding (1230). Remarkably, an ESV identified as *Photobacterium* sp. increased from <0.5% of amplicons in pre-feeding samples to 34% of

amplicons during this period (Figure 4.1B; Table C.2). Its abundance peaked at 1400, and then decreased sharply over the remaining time points. Other significantly varying ESVs that followed the same trend, but were less abundant, included unclassified members of the Vibrionaceae and diverse members of the phylum Firmicutes; roughly one-third of the significantly varying ESVs were identified as bacteria of the class Bacilli (Firmicutes), with all but one of these showing multi-fold increases in representation from the 1100 to 1230 time point (Table C.2). A smaller number of significantly varying ESVs showed the opposite trend, decreasing in representation after feeding. These included ESVs of the Firmicutes genus Streptococcus and unclassified members of the order gammaproteobacterial order Alteromonadales (Table C.2).

Of the 57 significantly varying ESVs, all but one were also detected in the food, albeit at relatively low proportional abundance (Table C.2). Food-associated ESVs included the intestinally abundant *Photobacterium* sp. ESV, which comprised ~2% of food sequences. In contrast, the most dominant food-associated sequences (>10%), belonging to the Cyanobacteria (Figure C.2), were not among the significantly varying ESVs in the intestine and contributed negligibly to the intestinal dataset.

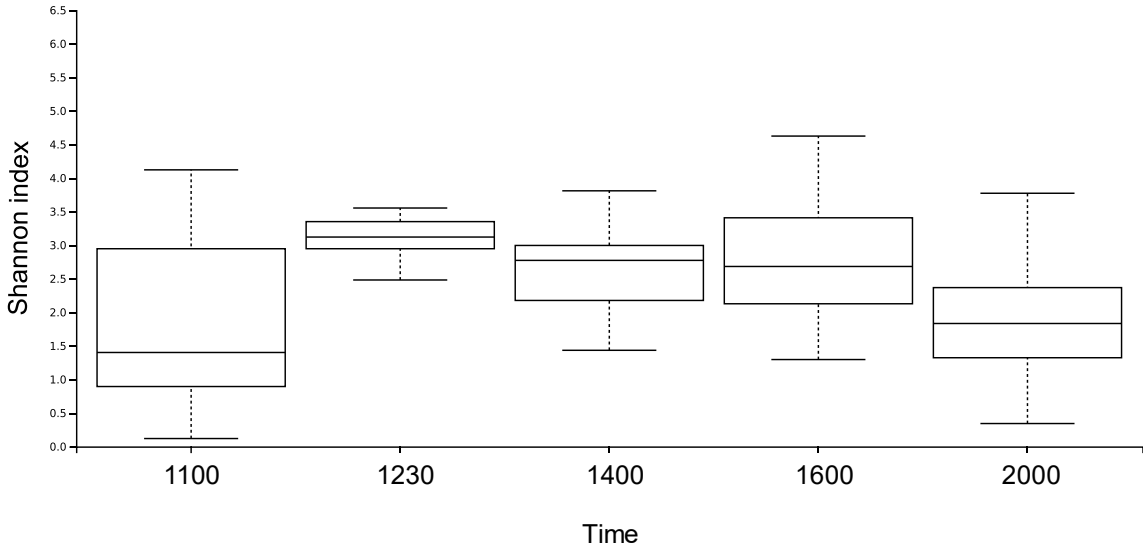

Figure 4.3. Average Shannon diversity varies according to time of sampling. Shannon diversity at 1230 was significantly higher compared to every other time point, excluding 1600; diversity at 2000 was significantly lower compared to every other time point, excluding 1100 (Kruskal Wallis pairwise, corrected p-value <0.05, n=17 fish at 1100 and 2000 and 18 fish at 1230, 1400, and 1600).

4.3.2 Predicted metagenome content

Metagenome prediction based on 16S rRNA gene profiles identified 329 gene categories. Of these, none were differentially represented among the three feeding frequency treatments, based on samples collected at the first 1100 time point of the diel-sampling period. In contrast, 111 functional gene categories were predicted to vary in abundance between unfed (1100 and 2000) and fed (1230, 1400, 1600) time points (DESEQ2, adjusted $p<0.05$, Table C.3). Functions predicted to be enriched in the fed state included those associated with bacteria-bacteria or bacteria-host interactions, including bacterial invasion of epithelial cells, infection by *Vibrio*, secretion, motility, and chemotaxis. In contrast, unfed time points were dominated by anabolic and catabolic functions (Figure

4.4, Table C.3). Of the top 50 predicted functional categories that were significantly enriched in the unfed time points, 76% (38/50) were classified broadly as metabolism, degradation, or biosynthesis, including those for amino acid metabolism, fatty acid metabolism, secondary bile acid production, sphingolipid biosynthesis, and the degradation of a wide range of organic compounds. In comparison, metabolism, degradation, or biosynthesis-associated categories represented only 18% of those enriched in the fed datasets.

4.3.3 Differential transcription

Metatranscriptome sequencing yielded 2,188,905 non-host, mRNA reads; per-sample counts ranged from 58,611 to 681,223 (Table C.4). Of these, 166,307 reads (across all samples) were classified as bacterial and had functional matches in the SEED database (a constantly updated repository for genomic sequence information; Overbeek et al. 2005), representing 1,063 pathways (second level of the SEED classification). Within these SEED pathways, a total of 267 genes showed differential transcription between fed and unfed samples ($p<0.05$, DESeq, Table C.5). Of the top 100 most differentially transcribed genes (based on adjusted p-value), 62 were at higher abundance in the unfed state (Figure 4.5). Consistent with the metagenome predictions, unfed transcriptomes were enriched in functions associated with metabolism, notably carbohydrate and amino acid metabolism. Over one third of the genes most enriched in the unfed transcriptomes were associated with diverse steps of carbohydrate utilization (compared to 16% of the fed-enriched genes), including several associated with pyruvate metabolism and the citric acid cycle (EC 1.2.4.2, EC 6.2.1.5, EC 4.2.1.3, EC 1.1.5.4, EC 2.3.1.12, EC 1.1.2.3, EC 2.3.3.5), fermentation or the metabolism of fermentation intermediates (EC 1.1.2.3, EC

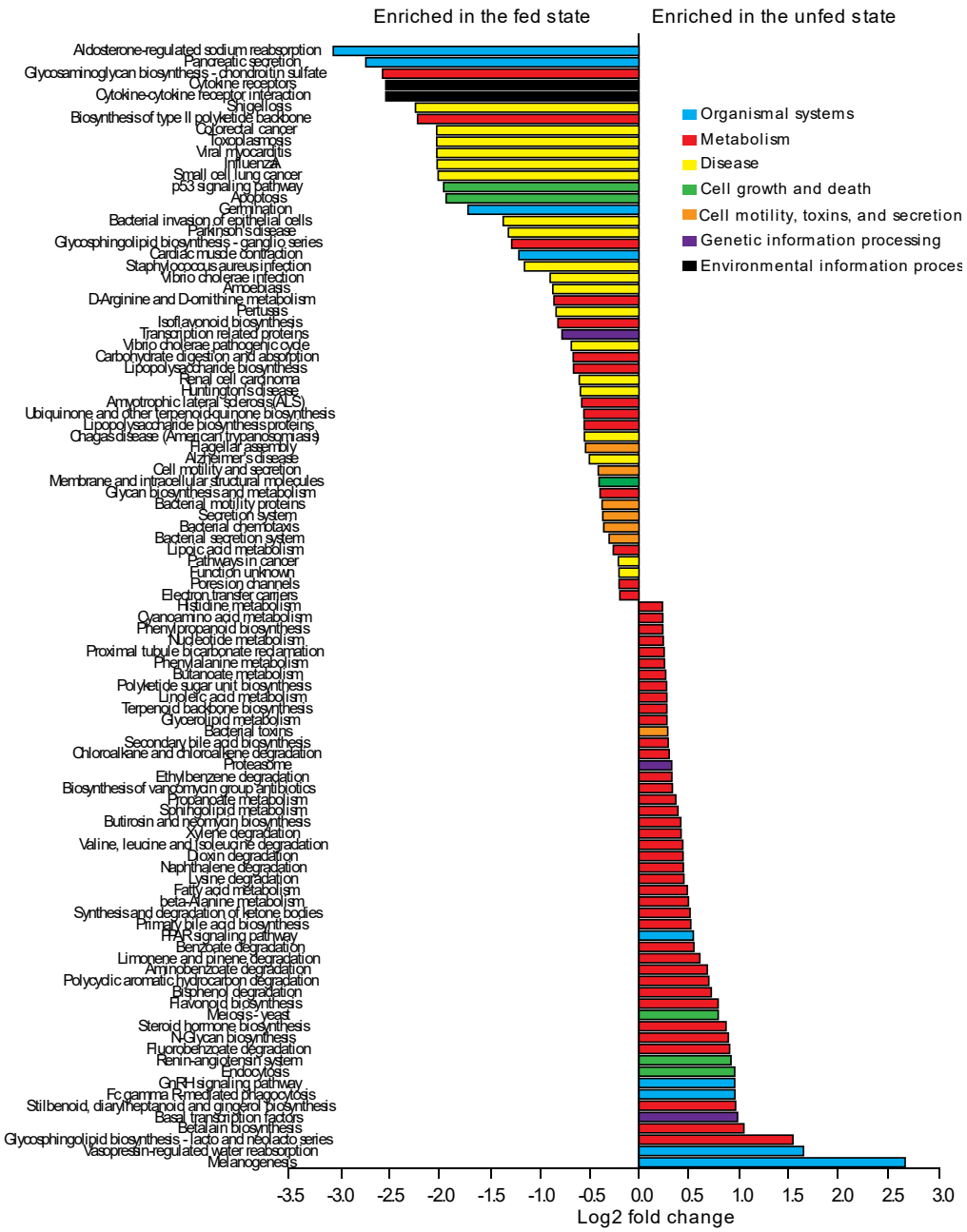

Figure 4.4. Predicted functional gene categories vary significantly between unfed (1100 and 2000) and fed (1230, 1400, 1600) time points with fed time points enriched in disease-associated pathways and unfed time points enriched in diverse metabolic pathways. Metagenomes were predicted from amplicon data using PICRUST and the 3rd hierarchical level of KEGG. Differential abundance was evaluated using DESEQ2 in R with gene categories having an adjusted p-value <0.05 shown here. Colors represent larger subsystem categories in KEGG. Only the top 99 of 111 significant pathways are plotted (n=17 fish at 1100 and 2000 and 18 fish at 1230, 1400, and 1600).

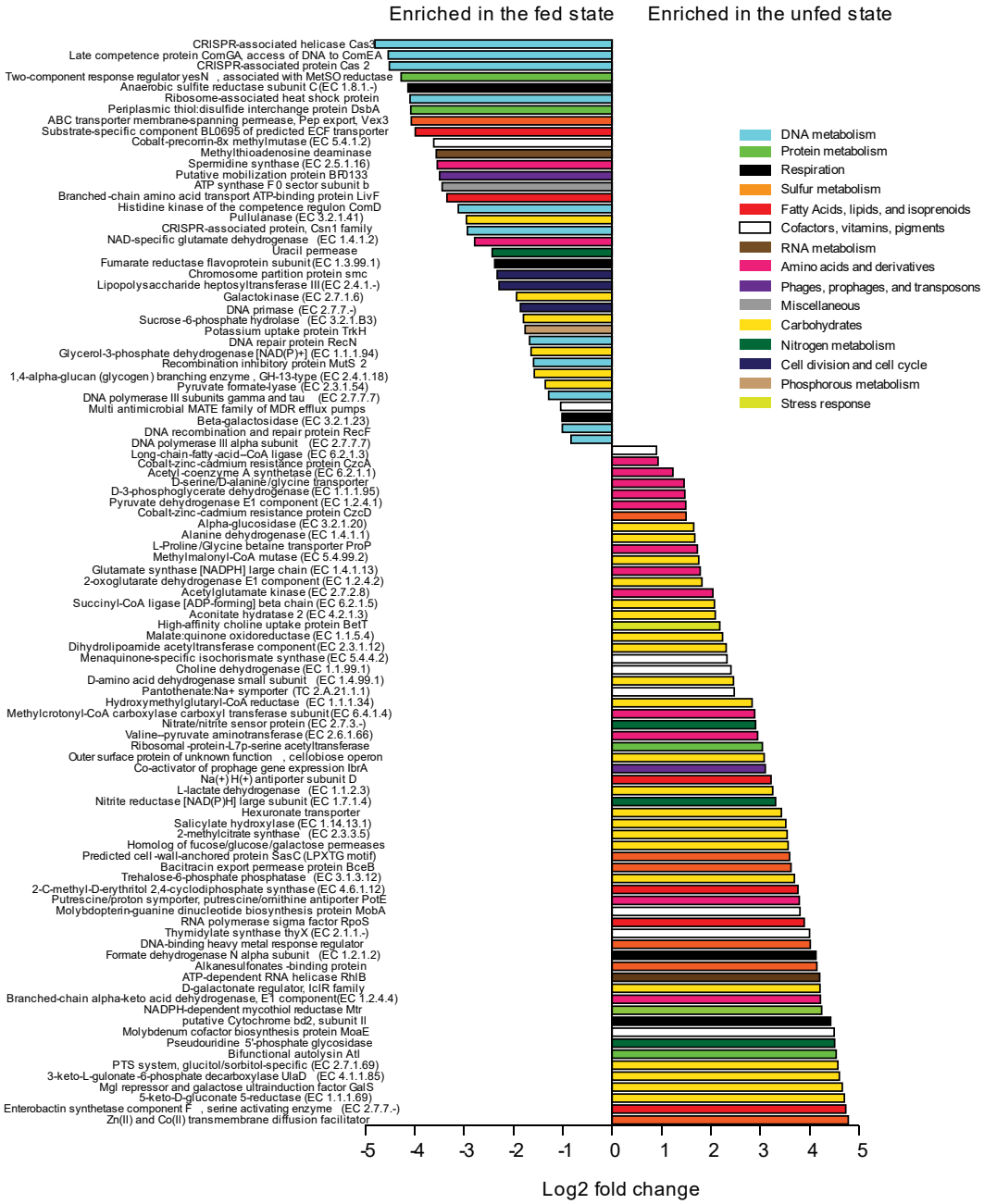

Figure 4.5. Gene expression varies significantly over a diurnal feeding cycle. Top SEED functional genes (by adjusted p-value) showing differential expression between fed and unfed states in transcriptomic data (1100 vs 1230 time points, n=5 for each). Differential abundance was tested with DESEQ2 in R and all functional categories shown vary significantly ($p<0.05$). Colors represent larger subsystem categories in SEED.

2.3.3.5), and the degradation of cellulose or other complex organic molecules (EC 1.14.13.1, EC 3.1.3.12, EC 1.1.1.69). In contrast, fed state transcriptomes were enriched in genes of DNA metabolism. This included several genes associated with CRISPR defense systems to cope with foreign DNA, as well as genetic elements associated with recombination (Figure 4.5).

4.4 Discussion

We used a popular and commonly bred marine aquarium species, the clownfish *Premnas biaculeatus*, as a model to explore how feeding events and feeding frequency alter the gut microbiome. Quantifying microbiome changes over the short time frame of a feeding and digestive cycle is critical for evaluating if microbiome studies should standardize sampling around feeding schedules, characterizing microbiome stability, distinguishing resident microbiome members from transient ones, and determining the extent to which the microbiome can be shaped by changes in feeding strategy independent of diet type.

In our experiment, gut microbiome composition and predicted metabolic function varied significantly over a 24-hour cycle (evaluated over two days). This cycle included one feeding event per day, and the most substantial changes in microbiome composition were evident when samples were grouped in relation to this event ("fed" vs "unfed"). Microbiome alpha diversity also spiked after feeding, suggesting that new microbes were introduced via food or that feeding changed the growth dynamics of resident microbes, or both. These patterns implicate feeding, rather than other host diel rhythms, as the primary driver of microbiome change over short (hourly) timescales in our experiment.

In contrast, microbiome communities, when analyzed at the start of the two-day diel sampling period did not group based on the frequency of feeding over the prior 30 days, suggesting that prior feeding frequency did not have a lasting restructuring effect on the microbial community.

Variation in microbiome composition over the diel period was driven largely by individual sequence variants, notably a member of the gammaproteobacterial genus *Photobacterium*. Its abundance in the intestine increased from near zero immediately before feeding to over one third of all sequences 1.5 hours post-feeding. This ESV was present in the food and therefore its post-feeding increase may be due to the arrival of new cells in the intestine, a process consistent with the observed post-feeding increase in alpha diversity. However, feeding also may have stimulated growth of cells already in the intestine. Indeed, *Photobacterium* species are common in fish microbiomes (Xing et al. 2013, Givens et al. 2015). Some species are pathogens, while others play mutualistic or commensal roles (Chabrillon et al. 2005, Urbanczyk et al. 2011). Members of the genus are typically facultatively aerobic chemoorganotrophs, motile via flagella, and employ diverse mechanisms for extracellular signaling and host interaction, including multiple virulence factors (Labella et al. 2017). *Photobacterium* genomes often contain a high number of rRNA operons (often >10; Rastogi et al. 2009), which biases estimates of the proportional abundance of this genus using 16S rRNA gene data. Nonetheless, the magnitude of change in *Photobacterium* sequence abundance after feeding is significant and implicates this taxon as an opportunistic member of the gut whose abundance and activity are closely linked to food availability. Indeed, the genus has been suggested to play a role in fish digestion, potentially by aiding the breakdown of chitin (Itoi et al.

2006, Ray et al. 2012, Egerton et al. 2018). Follow-up experiments that vary the chitin content of the diet could be used to test a linkage between *Photobacterium* population oscillations and chitin metabolism.

Other ESVs also fluctuated dramatically over the diel period but were less abundant (typically <0.5%). Like *Photobacterium*, the majority of these bacteria were also detected at low abundance in the food, barely detectable in pre-feeding samples, and spiked in representation immediately post-feeding (1230). Many of these taxa were members of the phylum Firmicutes, including diverse genera of lactic acid bacteria (LAB; order Lactobacillales) such as *Lactobacillus*, *Vagococcus*, *Leuconostoc*, and *Streptococcus* (Table C.2). LAB are common in vertebrate gut microbiomes, including of fishes in which *Lactobacillus* diversity in particular has been shown to be highly responsive to diet shifts (Desai et al. 2012, Schmidt et al. 2016). In mammalian systems, LAB have been shown to vary in abundance during feeding and non-feeding phases, although the nature of this variation was taxon-specific and varied depended on the timing (frequency) of feeding and also on diet type (Zarrinpar et al. 2014). LAB, and Firmicutes in general, are thought to be important to host carbohydrate metabolism through fermentation and have been associated with efficient dietary energy harvest (Turnbaugh et al. 2006, Guo et al. 2008, Krajmalnik-Brown et al. 2012); we hypothesize that LAB are likely playing a similar role in fermentation and energy extraction in the clownfish gut, although the specific dietary compounds supporting LAB catabolism in this system remain to be identified. Our findings corroborate prior evidence suggesting that LAB are among the most responsive to feeding, with many taxa showing a positive response, a factor that may also contribute to their role in energy extraction from food.

The post-feeding spike in abundance of *Photobacterium* and other diverse ESVs, coupled with their detection in the food, implicate microbial attachment to food as a major determinant of intestinal microbiome composition. However, other taxonomic groups that were much more abundant in the food (e.g., *Pseudoalteromonas*, diverse Cyanobacteria; Figure C.4) were not among those showing significant diurnal variation in the intestine microbiome. This suggests that the observed post-feeding spikes by certain intestinal microbes were not due exclusively to the influx of dead cells or DNA. Rather, microbes such as *Photobacterium* or LAB may be adept at both surviving passage through the stomach and exploiting the intestinal environment for growth, at least in the short term. It is also possible that these taxa are taking advantage of metabolites produced by other microbes or that the introduction of food provides them a competitive advantage by altering physical conditions or promoting growth-limiting factors like phages. Further work is needed to assess how the abundance and metabolism of these feeding-responsive taxa may vary with diet, including in omnivorous hosts such as *Premnas biaculeatus*, and to what extent these taxa persist in the intestine without regular input of new food-associated cells.

Other taxa, some of which were highly abundant, exhibited less dramatic fluctuation over the diel period and were not detected in the food. An ESV most closely related to the Firmicutes bacterium *Clostridium perfringens* dominated (>50%) the intestinal microbiomes in the unfed stages and also remained abundant after feeding. *Clostridium* bacteria are common in the intestine of fishes (Kim et al. 2007, Clements et al. 2007, Givens et al. 2015) and other vertebrates, including in humans where this genus has been associated with mucus scavenging (Tailford et al. 2015) and in mice where

Clostridium has been shown to vary over a daily cycle (Thaiss et al. 2014, Zarrinpar et al. 2014). These diverse fermenters are some of the first taxa to colonize human infants, are believed to localize to particular epithelial cells in the colon, and are important to colonic health by producing butyrate as an energy source for colonocytes (Lopetuso et al. 2013). In herbivorous fishes, *Clostridium* species have been associated with potentially beneficial roles in vitamin and fatty acid synthesis (Balcázar et al. 2006) and the production of metabolic enzymes for catabolism (Ramirez and Dixon 2003). Species such as *C. perfringens* are common food-born pathogens in humans and have been found previously in diverse fishes (e.g., Matches et al. 1974, Sabry et al. 2016). In our study, the proportional representation of *C. perfringens* was undoubtedly influenced by swings in taxa such as *Photobacterium* that increased rapidly after feeding (and vice versa). However, the overall high representation of *C. perfringens* at both fed and unfed time points suggests that this taxon may be physically associated with the intestinal lining, rather than the transitory stool. Association with the intestinal mucosal epithelium would be consistent with the mucolytic capabilities observed previously in *Clostridium* (Deplancke et al. 2002) and may suggest *Clostridium* as a comparatively persistent microbiome member across changes in diet or food availability.

Significant changes in community composition over the feeding cycle coincided with differences in predicted gene content and gene transcription in the fed and unfed states. The results were generally consistent between analyses, with time points immediately after feeding enriched in pathways involved in bacterial secretion systems, pathogen interaction with hosts, cell motility, and coping with foreign DNA (e.g., CRISPR). In contrast, unfed time points were enriched in transcripts and predicted genes

of diverse metabolic processes, particularly those involved in the catabolism of diverse organic substrates, including through fermentation, suggesting these periods as important for microbial degradation of dietary compounds. This enrichment of genes of fermentative carbohydrate and amino acid metabolism is likely linked to *Clostridia*, the dominant taxonomic group in the unfed state (Figure 4.1B) and one known to play diverse fermentative roles in other gut systems. These patterns are undoubtedly influenced by the dramatic fluctuations in the feeding-responsive members of the community, primarily *Photobacterium*, but likely also the unclassified members of the Vibrionaceae that increased during fed time points (Table C.2). Indeed, comparative analysis of *Photobacterium* genomes has revealed high numbers of CRISPR arrays, prophage sequences, and genomic islands, suggesting that phage infection and gene mobilization may be common in this genus (Machado and Gram 2017); other Vibrionaceae genomes share similar features (Gu et al. 2009, Lin et al. 2018).

These results suggest that the period shortly after feeding may be a time of increased bacteria-bacteria and bacteria-host interaction. The movement of food into the intestine may stimulate bacteria living in association with the host mucus layer to mobilize and attach to food particles. These early responders therefore may be those cells best equipped to navigate a dynamic and spatially structured environment. Some of these are likely already resident in the intestine, although perhaps at low proportional abundance. However, our results suggest that feeding also serves as the major inoculation event through which new cells enter the system and potentially try to establish residence, a process that would presumably also be characterized by social interactions, attachment, and motility.

4.5 Conclusions

These results confirm that feeding is a major restructuring force in intestinal microbiomes over a short timeframe (hours). This restructuring involves swings in proportional abundance that differ among microbial types, likely due to differences in metabolic and spatial niche (for example, attachment to food versus residence in an intestinal epithelial biofilm), and potentially also interactions among neighboring microbes. The patterns reported here identify taxa to target for comparisons of how opportunistic, feeding-responsive microbes of the intestine differ ecologically from more persistent, and potentially commensal, members. The large post-feeding changes involving food-associated microbes indicate high connectivity between external and intestinal microbial pools. Animals rarely if ever consume sterile food, even in captivity, and also exhibit significant variation in feeding schedule. This variation should be addressed in comparative microbiome studies, particularly those involving wild animals or small numbers of replicates, but also those focused on model systems (e.g., gnotobiotic mice). Indeed, a growing body of evidence, including from this study, suggests that the vertebrate gut microbiome can exhibit significant short-term fluctuation. Sampling a microbiome at different points relative to the last feeding event will therefore likely yield different conclusions about microbiome composition and function.

4.6 Methods

4.6.1 Feeding experiments

A single cohort of 120, six-month old maroon clownfish (*Premnas biaculeatus*) was obtained from Sustainable Aquatics (Jefferson City, TN) and allowed to acclimate for two weeks in an artificial seawater system at Georgia Tech. All fish were fed a 0.8

mm dry pellet composite of krill meal, fish meal, squid meal, wheat gluten, potato starch, fish oil, spirulina, astaxanthin, and garlic oil produced by Sustainable Aquatics. During acclimation, fish were fed once daily at 1100. Following acclimation, fish were equally divided into 12 identical, 10-gallon tanks, all of which were connected to the same recirculating water system. Individual tanks were randomly assigned to one of three treatment groups based on feeding frequency: a 4x treatment with feeding of 2.1 mg of food per fish 4 times daily at 0800, 1100, 1600, and 2000; a 1x treatment with feeding of 8.4 mg of food per fish once daily at 1100; and a 0.33x treatment with feeding of 25 mg of food per fish once every 3 days at 1100. While we could not verify the actual amount of food eaten per fish, the amount of food applied per feeding in the 0.33x treatment was consumed entirely (the food was buoyant, allowing us to track consumption), indicating that this amount did not exceed the maximum clearance rate of each fish group. The feeding treatments were administered for 30 days and water quality was monitored daily. Samples of the tank water microbiome were collected from each treatment at the midpoint and end of the experiment by filtering 1 L of water onto 0.2 μm 25 mm disc filters (n=3 per time point per treatment). Samples (n=3) to analyze microbes in the food were collected at the end of the experiment. A subset of fish (n=6) were sacrificed before initiating feeding treatments and at day 15 (one from each treatment) during the experiment by submerging individuals in a sterile water bath containing MS-222. After euthanization, a ventral cut was made on each fish to expose the gut cavity to preservative, and each fish was then preserved in RNA/DNA stabilizing buffer (25 mM sodium citrate, 10 mM EDTA, 5.3M Ammonium sulfate, pH 5.2) and stored frozen. After 30 days, the remaining fish from each treatment group were sacrificed at various

daily time points over a two day period (3 treatment groups X 3 replicates per daily time point X 2 days of sampling each time point = 18 total individuals sacrificed per time point): immediately pre-feeding (1100, n=17), 1.5 hours post-feeding (1230, n=18), three hours post-feeding (1400, n=18), 5 hours post-feeding (1600, n=17), and nine hours post-feeding (2000, n=17). (For some time points, we obtained data from 17 rather than 18 individuals due to sample/fish loss) During this two-day sacrifice period, all fish were fed only once per day at 1100 regardless of original feeding frequency treatment. Fish were euthanized and preserved as described above. Prior to DNA/RNA extraction, whole intestines were removed from each fish via sterile dissection.

4.6.2 DNA/RNA extractions and amplicon generation

Total DNA and RNA was extracted from intestinal contents and water/food samples using the Mobio PowerMicrobiome® RNA Isolation kit. Total extracts were split into DNA and RNA pools and treated with either DNase or RNase respectively. Before beginning the kit protocol for fish samples, a longitudinal cut was made along the length of the previously dissected intestine. The intestine and associated contents were placed inside a bead-beating tube provided with the kit, vortexed for 5-10 seconds, and the remaining intact intestine was removed and discarded to minimize host signal. The remainder of the extraction followed standard procedures for the Mobio PowerMicrobiome® RNA Isolation kit.

Illumina sequencing of 16S rRNA gene amplicons was used to assess microbiome community composition in a subset of experimental fish (90 fish total: 6 per time point * 5 time points * 3 treatments). Amplicons were synthesized using Platinum® PCR SuperMix (Life Technologies) with primers F515 and R806 spanning the V3-V4 region

of the 16S rRNA gene (Caporaso et al., 2011). Forward and reverse primers were modified to include Illumina sequencing adapters according to Kozich et al. (2013) and barcoded by sample to maintain integrity of biological replicates. Approximately 5 ng of starting DNA was used as template for each PCR reaction. Negative controls using sterile water were included with each set of PCR reactions. Amplification was performed using denaturation at 94°C (3 min), followed by 30 cycles of denaturation at 94°C (45 sec), primer annealing at 55°C (45 sec), primer extension at 72°C (90 sec), and a final extension at 72°C for 10 min. Amplicons were verified using gel electrophoresis, purified using Diffinity RapidTip2 PCR purification tips (Diffinity Genomics, NY), and quantitated fluorometrically using the Qubit (Life Technologies). Barcoded amplicons from all samples were pooled at equimolar concentrations and sequenced on an Illumina MiSeq using a 500 cycle kit with 10% PhiX added to increase sequence diversity.

4.6.3 Amplicon sequence analysis

Amplicon sequence data were sorted by sample according to barcode, quality-controlled (removed bases <Q30 and sequences less than 150 base pairs), and clustered into exact sequence variants (ESVs) using Quantitative Insights Into Microbial Ecology (QIIME2, version 2018.2) with plugins demux and deblur (Amir et al. 2017). Sequences were summarized using the feature-table function and aligned, and representative sequences from each cluster were arranged in a phylogenetic tree using FastTree. The feature table with water and food samples included was rarified to 10,321 sequences per sample and alpha and beta diversity explored using the q2-diversity plugin. Alpha diversity among sample groupings was compared using the Shannon Diversity metric, and beta diversity was compared using Bray-Curtis distance. A second feature table was

created to include only experimental fish samples and was rarefied to 20,562 sequences per sample. This table excluded 1 sample from the 1100 time point and one sample from the 2000 time point due to low sequence yield. Taxonomy was assigned to sequence clusters using a pre-trained Naïve Bayes classifier and the q2-feature-classifier plugin. This classifier was trained on the Greengenes 13_8 99% OTUs database which included a 250 base pair segment from the V4 region of the 16S. Differentially abundant taxa (grouped at the genus level) between sample groupings (feeding frequency, time of sampling) were identified using ANCOM (Mandal et al. 2015). The ability to predict sample groupings based on taxonomic composition was assessed using the q2-sample-classifier function and a random forest model for both feeding frequency and time of sampling.

Predicted metagenomes were constructed from amplicon data using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt, Langille et al. 2013). This analysis was based on an OTU table generated using QIIME1 with the pick_closed_reference_otus.py command, OTUs clustered at 97% sequence similarity, and Greengenes version 13_8. QIIME1 was used at this step since the output OTU table is compatible with PICRUSt. The OTU table was normalized by dividing each OTU by the known/predicted 16S rRNA gene copy number with the normalize_by_copy_number.py command in PICRUSt. Functional predictions based on KEGG categories were made via the predict_metagenomes.py command and collapsed at the 3rd hierarchical level using the categorize_by_function.py command. Predicted functional categories that differed in abundance between feeding treatments and sampling time points were identified using DESeq2 in R.

For ease of comparison and because taxonomic analysis showed high similarity among post-feeding time points, samples were grouped into “fed” (1230, 1400, and 1600) and “unfed” (1100 and 2000) categories. The “unfed” grouping is presumed to include time points at which most of the ingested food has left the stomach; however, it is estimated to take ~36 hours for food to completely pass through the clownfish intestine (Ling and Ghaffar, 2014). Therefore, it is likely that all samples contained some amount of food regardless of sampling time.

4.6.4 RNA sequencing and analysis

The amplicon-based analyses (above) suggested a large difference in composition and predicted microbiome function between samples in the fed and unfed states, primarily at the time points before and shortly after feeding. We therefore chose samples on day 1 from the fed (1230 time point, n=5) and unfed (1100 time point, n=5) state as focal points to analyze transcription of metabolic genes in response to feeding.To generate mRNA data, DNA was first removed from an aliquot of each RNA extract using the TURBO DNA-*free*™ kit (Invitrogen). We next attempted to enrich samples for non-host, non-ribosomal RNA using the MICROBEnrich™ kit (Ambion) and the Ribo-Zero rRNA removal kit (Illumina) following recommended procedures. Barcoded cDNA libraries were prepared from microbial mRNA-enriched RNA using the ScriptSeq kit by Illumina and sequenced (250X250 bp) on one lane of an Illumina HiSeq 2500 flowcell in Rapid mode.

Using Trim Galore! (Krueger, 2015), FastQ sequence files were trimmed to remove bases with quality scores less than Q30 and to discard sequences with fewer than 75 bp. To identify host-like RNA, trimmed sequences were mapped against a reference

genome of *Amphiprion ocellaris* (the only publicly available clownfish genome, NCBI accession PRJNA407816) using default parameters in BBMap. The unmapped reads were retained and ribosomal reads removed using riboPicker (Schmieder et al. 2011). Unmapped, non-ribosomal reads were used as input for BLASTX queries against the NCBI nr database (October, 2017 release) via stand-alone BLAST version 2.6.0+. BLASTX results with bit score >50 were imported into MEGAN6 (Huson et al. 2016) with taxonomy assigned to reads using NCBI taxonomy and MEGAN6's LCA algorithm. Functional annotation was performed in MEGAN using the SEED database. SEED is a composite, regularly curated genome annotation database (Overbeek et al. 2005). Gene categories that differed significantly in abundance between "fed" and "unfed" transcriptomes were identified using DESEQ2 in R, based on the raw gene count matrix exported from MEGAN6.

4.7 Acknowledgements

We thank John and Matthew Carberry at Sustainable Aquatics for stimulating discussion about marine aquaculture, and for their deep knowledge of clownfish husbandry in general.

4.8 References

Albenberg, L., Esipova, T. V., Judge, C. P., Bittinger, K., Chen, J., Laughlin, A., ... & Thom, S. R. 2014. Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. *Gastroenterology 147*(5):1055-1063.

Amir, A., McDonald, D., Navas-Molina, J. A., Kopylova, E., Morton, J. T., Xu, Z. Z., ... & Knight, R. 2017. Deblur rapidly resolves single-nucleotide community sequence patterns. *MSystems*, *2*(2):e00191-16.

Balcázar, J. L., De Blas, I., Ruiz-Zarzuela, I., Cunningham, D., Vendrell, D., and Muzquiz, J. L. 2006. The role of probiotics in aquaculture. *Vet. Microbiol.* 114:173–186. doi: 10.1016/j.vetmic.2006.01.009

Belenguer, A., Duncan, S. H., Calder, A. G., Holtrop, G., Louis, P., Lobley, G. E., & Flint, H. J.. 2006. Two routes of metabolic cross-feeding between Bifidobacterium adolescentis and butyrate-producing anaerobes from the human gut. *Applied and environmental microbiology 72*(5):3593-3599.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., ... & Knight, R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, 2011;*108*(Supplement 1):4516-4522.

Comstock, L. E., & Coyne, M. J. Bacteroides thetaiotaomicron: a dynamic, niche-adapted human symbiont. *Bioessays*, 2003;*25*(10):926-929.

Conlon MA, Bird AR. The impact of diet and lifestyle on gut microbiota and human health. *Nutrients*. 2014;7(1):17-44.

Chung, H., Pamp, S. J., Hill, J. A., Surana, N. K., Edelman, S. M., Troy, E. B., ... & Umesaki, Y. Gut immune maturation depends on colonization with a host-specific microbiota. *Cell*, 2012;*149*(7):1578-1593.

Clements, K. D., Pasch, I. B., Moran, D., and Turner, S. J. Clostridia dominate 16S rRNA gene libraries prepared from the hindgut of temperate marine herbivorous fishes. *Mar. Biol.* 2007;150:1431–1440. doi: 10.1007/s00227-006-0443-9

Crawford, P. A., Crowley, J. R., Sambandam, N., Muegge, B. D., Costello, E. K., Hamady, M., ... & Gordon, J. I. Regulation of myocardial ketone body metabolism by the gut microbiota during nutrient deprivation. *Proceedings of the National Academy of Sciences*, 2009;*106*(27):11276-11281.

Deplancke B, Vidal O, Ganessunker D, Donovan SM, Mackie RI, Gaskins HR. Selective growth of mucolytic bacteria including Clostridium perfringens in a neonatal piglet model of total parenteral nutrition. *Am J Clin Nutr*. 2002 Nov;76(5):1117-25.

Desai, A.R., Links, M.G., Collins, S.A., Mansfield, G.S., Drew, M.D., Van Kessel, A.G.and Hill, J.E. Effects of plant-based diets on the distal gut microbiome of rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 2012;350**:**134–142.

De Vuyst, L., & Leroy, F. Cross-feeding between bifidobacteria and butyrate-producing colon bacteria explains bifdobacterial competitiveness, butyrate production, and gas production. *International journal of food microbiology*, 2011;*149*(1):73-80.

Donaldson, G. P., Lee, S. M., & Mazmanian, S. K. Gut biogeography of the

bacterial microbiota. *Nature Reviews Microbiology*, 2016;*14*(1):20.
Egerton, S., Culloty, S., Whooley, J., Stanton, C., & Ross, R. P. The gut microbiota of marine fish. *Frontiers in Microbiology*, 2018;*9*.
Espey, M. G. Role of oxygen gradients in shaping redox relationships between the human intestine and its microbiota. *Free Radical Biology and Medicine*, 2013;*55*:130-140.
Galland L. The Gut Microbiome and the Brain. *Journal of Medicinal Food.* 2014;17(12):1261-1272.
Gerritse, J., Schut, F., & Gottschal, J. C. Mixed chemostat cultures of obligately aerobic and fermentative or methanogenic bacteria grown under oxygen-limiting conditions. *FEMS Microbiology Letters*, 1990;*66*(1-3):87-93.
Gilbert JA, Quinn RA, Debelius J, Xu ZZ, Morton J, Garg N, et al. Microbiome-wide association studies link dynamic microbial consortia to disease. *Nature*. 2016;535:94.
Givens, C. E., Ransom, B., Bano, N., & Hollibaugh, J. T. Comparison of the gut microbiomes of 12 bony fish and 3 shark species. *Marine Ecology Progress Series*, 2015;*518*:209-223.
Gross, E. M., Brune, A., & Walenciak, O. Gut pH, redox conditions and oxygen levels in an aquatic caterpillar: potential effects on the fate of ingested tannins. *Journal of Insect Physiology*, 2008;*54*(2):462-471.
Gu J, Neary J, Cai H, et al. Genomic and systems evolution in *Vibrionaceae* species. *BMC Genomics*. 2009;10(Suppl 1):S11. doi:10.1186/1471-2164-10-S1-S11.
Hehemann, J. H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., & Michel, G. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature*, 2010;*464*(7290):908.
Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., ... & Tappu, R. MEGAN community edition-interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS computational biology*, 2016;*12*(6):e1004957.
Itoi, S., Okamura, T., Koyama, Y., and Sugita, H. Chitinolytic bacteria in the intestinal tract of Japanese coastal fishes. *Can. J. Microbiol.* 2006;52:1158–1163. doi: 10.1139/w06-082
Johnson, K. S., & Barbehenn, R. V. Oxygen levels in the gut lumens of herbivorous insects. *Journal of Insect Physiology*, 2000;*46*(6):897-903.
Kaczmarek, J. L., Thompson, S. V., & Holscher, H. D. Complex interactions of circadian rhythms, eating behaviors, and the gastrointestinal microbiota and their potential impact on health. *Nutrition reviews*, 2017;*75*(9):673-682.
Kim, D. H., Brunt, J., & Austin, B. Microbial diversity of intestinal contents and mucus in rainbow trout (Oncorhynchus mykiss). *Journal of Applied Microbiology*, 2007;*102*(6):1654-1664.
Kohl, K. D., Amaya, J., Passement, C. A., Dearing, M. D., & McCue, M. D. Unique and shared responses of the gut microbiota to prolonged fasting: a comparative study across five classes of vertebrate hosts. *FEMS microbiology ecology*, 2014;*90*(3):883-894.
Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D.

Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and environmental microbiology*, 2013;*79*(17):5112-5120.

Krajmalnik-Brown R, Ilhan ZE, Kang DW, DiBaise JK. Effects of gut microbes on nutrient absorption and energy regulation. *Nutr. Clin. Pract*. 2012;27:201–14.

Labella, A. M., Arahal, D. R., Castro, D., Lemos, M. L., & Borrego, J. J. Revisiting the genus Photobacterium: taxonomy, ecology and pathogenesis. *Int. Microbiol*, 2017;20:1-10.

Langille, M. G., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., ... & Beiko, R. G. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature biotechnology*, 2013;*31*(9): 814.

Lee, W. J., & Hase, K. Gut microbiota–generated metabolites in animal health and disease. *Nature chemical biology*, 2014;*10*(6):416.

Li, F., Jiang, C., Krausz, K. W., Li, Y., Albert, I., Hao, H., ... & Gonzalez, F. J.. Microbiome remodelling leads to inhibition of intestinal farnesoid X receptor signalling and decreased obesity. *Nature communications*, 2013;*4*:2384.

Li, G., Xie, C., Lu, S., Nichols, R. G., Tian, Y., Li, L., ... & Krausz, K. W. (2017). Intermittent fasting promotes white adipose browning and decreases obesity by shaping the gut microbiota. *Cell metabolism*,2017;*26*(4):672-685.

Lin H, Yu M, Wang X, Zhang X-H. Comparative genomic analysis reveals the evolution and environmental adaptation strategies of vibrios. *BMC Genomics*. 2018;19:135. doi:10.1186/s12864-018-4531-2.

Ling, K. M., & Ghaffar, M. A. Estimation of gastric emptying time (GET) in clownfish (Amphiprion ocellaris) using X-radiography technique. In *AIP Conference Proceedings* 2014;1614(1):624-628.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., & Henrissat, B.. The carbohydrate-active enzymes database (CAZy). *Nucleic acids research*, 2013;*42*(D1):D490-D495.

Lopetuso, L. R., Scaldaferri, F., Petito, V., & Gasbarrini, A. (2013). Commensal Clostridia: leading players in the maintenance of gut homeostasis. *Gut pathogens*, *5*(1), 23.

Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. *Nature*. 2012;489(7415):220-30.

Machado H, Gram L. Comparative Genomics Reveals High Genomic Diversity in the Genus Photobacterium. *Frontiers in Microbiology*. 2017;8:1204. doi:10.3389/fmicb.2017.01204.

Mandal, S., Van Treuren, W., White, R. A., Eggesbø, M., Knight, R., & Peddada, S. D.. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microbial ecology in health and disease*, 2015;*26*(1): 27663.

Matches J R, Liston J, Curran D. *Clostridium perfringens* in the environment. *Appl Microbiol*. 1974;28:655–660.

McNulty, N. P., Yatsunenko, T., Hsiao, A., Faith, J. J., Muegge, B. D., Goodman, A. L., ... & Chervaux, C.. The impact of a consortium of fermented milk strains on the gut microbiome of gnotobiotic mice and monozygotic twins. *Science translational medicine*, 2011;*3*(106):106ra106-106ra106.

Overbeek, R.F., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H.Y., Cohoon, M., de Crecy-Lagard, V., Diaz, N., Disz, T., Edwards, R., et al. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res,*. 2005;33:5691–5702.

Ramirez, R. F., and Dixon, B. A. Enzyme production by obligate intestinal anaerobic bacteria isolated from oscars (*Astronotus ocellatus*), angelfish (*Pterophyllum scalare*) and southern flounder (*Paralichthys lethostigma*). *Aquaculture* 2003;227:417–426. doi: 10.1016/S0044-8486(03)00520-9

Rastogi, R., Wu, M., DasGupta, I., & Fox, G. E. Visualization of ribosomal RNA operon copy number distribution. *BMC microbiology*, 2009;*9*(1):208.

Ray, A. K., Ghosh, K., and Ringø, E. Enzyme-producing bacteria isolated from fish gut: a review. *Aquac. Nutr.* 2012;18:465–492. doi: 10.1111/j.1365-2095.2012.00943.x

Sabry M, Abd El-Moein K, Hamza E, Abdel Kader F. Occurrence of *Clostridium perfringens* Types A, E, and C in Fresh Fish and Its Public Health Significance. *J Food Prot*. 2016;79(6):994-1000. doi: 10.4315/0362-028X.JFP-15-569. PubMed PMID: 27296604.

Schmieder, R., Lim, Y. W., & Edwards, R.. Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics*, 2011;*28*(3):433-435.

Schmidt, V., Amaral-Zettler, L., Davidson, J., Summerfelt, S. and Good, C. Influence of fishmeal-free diets on microbial communities in Atlantic Salmon (*Salmo salar*) recirculation aquaculture systems. *Appl Environ Microbiol* 2016;82:4470–4481.

Secor, S. M., & Carey, H. V.. Integrative physiology of fasting. *Comprehensive Physiology*, 2016.

Sung, J., Kim, S., Cabatbat, J. J. T., Jang, S., Jin, Y. S., Jung, G. Y., ... & Kim, P. J. Global metabolic interaction network of the human gut microbiota for context-specific community-scale analysis. *Nature communications*, 2017;*8*: 15393.

Tailford, L. E., Owen, C. D., Walshaw, J., Crost, E. H., Hardy-Goddard, J., Le Gall, G., ... & Juge, N. Discovery of intramolecular trans-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation. *Nature communications*, 2015;*6*:7624.

Thaiss, C. A., Zeevi, D., Levy, M., Zilberman-Schapira, G., Suez, J., Tengeler, A. C., ... & Kuperman, Y. Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. *Cell*, 2014;*159*(3):514-529.

Thaiss, C. A., Levy, M., Korem, T., Dohnalová, L., Shapiro, H., Jaitin, D. A., ... & Tuganbaev, T. Microbiota diurnal rhythmicity programs host transcriptome oscillations. *Cell*, 2016;*167*(6):1495-1510.

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

Urbanczyk, H., Ast, J. C., & Dunlap, P. V. Phylogeny, genomics, and symbiosis of Photobacterium. *FEMS microbiology reviews*, 2011;*35*(2):324-342.

Van Hoek, M. J., & Merks, R. M. Emergence of microbial diversity due to

cross-feeding interactions in a spatial model of gut microbial metabolism. *BMC systems biology*, 2017;*11*(1):56.

Voigt, R. M., Forsyth, C. B., Green, S. J., Engen, P. A., & Keshavarzian, A. Circadian rhythm and the gut microbiome. In *International review of neurobiology* 2016;131:193-205. Academic Press.

Xing, M., Hou, Z., Yuan, J., Liu, Y., Qu, Y., & Liu, B. Taxonomic and functional metagenomic profiling of gastrointestinal tract microbiome of the farmed adult turbot (Scophthalmus maximus). *FEMS microbiology ecology*, 2013;*86*(3):432-443.

Zarrinpar, A., Chaix, A., Yooseph, S., & Panda, S. Diet and feeding pattern affect the diurnal dynamics of the gut microbiome. *Cell metabolism*, 2014;*20*(6):1006-1017.

CHAPTER 5
CONCLUSIONS AND SUGGESTIONS

The contribution of host-associated microbiomes to host health and ecology is now widely recognized and understanding the factors regulating microbiome diversity and function is a central question in biology. However, the vast majority of our present knowledge on microbiome ecology is based on research in mammal systems and we remain naïve in our understanding of how the importance of different microbiome assembly processes might vary among diverse hosts. This is especially true for the most phylogenetically and ecologically diverse of the vertebrate groups, teleost fishes. In this dissertation, I used a combination of 16s rRNA gene sequencing, metatranscriptomics, and genomics to explore the effects of life stage transition and feeding on the gut microbiome of marine fish. Results identify key taxanomic and functional signatures of the marine fish microbiome that may serve as a starting point for future studies investigating the linkages between fish gut microbiomes and host health and ecology.

5.1 Microbiome changes across development

Results from Chapter 3 describe a diverse gut microbiome in two abundant reef fish families (Pomacentridae and Apogonidae) and highlight the importance of life stage in structuring microbiome composition. Although changes associated with development have been characterized in laboratory fish, the work in this chapter is one of the first studies to describe changes in the gut microbiome associated with an ecologically relevant life stage transition in natural populations of marine fish. A trend toward greater microbiome diversity in settled (older) individuals was observed following the transition from a pelagic larval phase to reef settlement. This trend is potentially explained by the

hypothesis that pre-settlement (larval) fish of diverse species are more similar to each other physiologically compared to adults, occupy the same niche (pelagic zone), and acquire gut microbes from a common environmental pool, perhaps due to similarities in diet. These factors could homogenize the larval microbiome across diverse host species, with microbiomes diversifying after fish settle on the reef and transition to adult feeding roles. However, this pattern is not uniform across host species, suggesting a need for discretion in concluding general trends in microbiome succession over development. Targeted studies which robustly sample a single host species across life stages in the wild would help us confirm these diversity patterns and understand exactly how host ecology (feeding and habitat changes) might contribute to these bulk changes in microbiome diversity.

Chapter 3 also identifies gut microbial taxa varying significantly in abundance between pre and post-settlement reef fish including bacteria from the families Endozoicomonaceae and Shewanellaceae which were more common in pre-settlement fish and Pasteurellaceae and Chromatiales enriched in post-settlement hosts. Differences in microbiome assembly and therefore composition between pre and post-settlement hosts are likely driven by a combination of factors, including physiological changes associated with development, changes in diet and potentially feeding frequency, and varying connectivity with microbes from the external environment. More research is needed to disentangle the relative contributions of these determinants over the settlement transition. However, identifying taxa that vary in abundance over development serves as a starting point for understanding how different microbiome assembly processes operate at various life stages.

Several individual bacterial OTUs were common (>70%) in both pre and post-settlement fish including OTUs identified to Endozoicomonaceae, Chromatiales, Vibrionaceae, *Shewanella*, *Photobacterium*, and *Salinivibrio costicola* (Ch. 3). This set represents a relatively abundant (35-65% of total 16S sequences) core microbiome across damselfish and cardinalfish, including potentially pathogenic or beneficial bacteria. Persistence over development may be an important indicator of commensal or mutualistic taxa ubiquitous to reef fish as they colonize the digestive tract early and remain a part of the system regardless of diet, developmental, or habitat changes. These taxa may be particularly useful to explore if interested in questions related to host development and health.

5.2 Feeding effects on the gut microbiome

Chapter 4 examined the effect of feeding on the gut microbiome of marine fish by sampling maroon clownfish (*Premnas biaculeatus*) siblings over various time points around a feeding cycle. Data confirmed feeding as a major restructuring force in intestinal microbiomes over a short timeframe (hours). While studies have shown significant effects of discrete diet types on the gut microbiome, the work in this chapter is among the first to explore these "diet independent" effects.

Restructuring of the gut microbiome over short time scales in clownfish involved swings in proportional abundance that differed among microbial types, likely due to differences in metabolic and spatial niche (for example, attachment to food versus residence in an intestinal epithelial biofilm), and potentially also interactions among neighboring microbes. Two individual ESVs identified to *Photobacterium* and *Clostridium perfringens* were particularly important in the guts of clownfish with these

two taxa alone comprising up to 70% of the microbiome signal at certain time points around feeding (Ch. 4).

Photobacterium (and many other ESVs at lower abundance) increased significantly following feeding indicating microbial attachment to food as a major determinant of intestinal microbiome composition. However, taxonomic groups that were much more abundant in the food (e.g., *Pseudoalteromonas*, diverse Cyanobacteria; Chapter 4) were not among those showing significant diurnal variation in the intestine microbiome. This suggests that the observed post-feeding spikes by certain intestinal microbes were not due exclusively to the influx of dead cells or DNA. Rather, microbes such as *Photobacterium* may be adept at both surviving passage through the stomach and exploiting the intestinal environment for growth, at least in the short term. It is also possible that these taxa are taking advantage of metabolites produced by other microbes or that the introduction of food provides them a competitive advantage by altering physical conditions or promoting growth-limiting factors like phages. Further work is needed to assess how the abundance and metabolism of these feeding-responsive taxa may vary with diet, including in omnivorous hosts such as *Premnas biaculeatus*, and to what extent these taxa persist in the intestine without regular input of new food-associated cells.

In contrast to the *Photobacterium* trend, an ESV most closely related to the Firmicutes bacterium *Clostridium perfringens* dominated (>50%) the intestinal microbiomes in the unfed stages and remained abundant after feeding. Clostridium species are fermenters of diverse compounds, are believed to localize to particular epithelial cells in the colon, and are important to colonic health by producing butyrate as

an energy source for colonocytes in mammals (Lopetuso et al. 2013). In herbivorous fishes, *Clostridium* species have been reported to have beneficial roles in vitamin and fatty acid synthesis (Balcázar et al. 2006) and the production of metabolic enzymes for catabolism (Ramirez and Dixon 2003). The overall high representation of *C. perfringens* at both fed and unfed time points suggests that this taxon may be physically associated with the intestinal lining, rather than the transitory stool in clownfish and potentially represents a mutualistic member of the fish microbiomes. An examination of the relationship between host health metrics (growth rate, disease resistance, etc.) and *Clostridium* abundance is warranted.

Significant changes in community composition over the clownfish feeding cycle coincided with differences in predicted gene content and gene transcription in the fed and unfed states. The results were generally consistent between analyses, with time points immediately after feeding enriched in pathways involved in bacterial secretion systems, pathogen interaction with hosts, cell motility, and coping with foreign DNA (e.g., CRISPR). In contrast, unfed time points were enriched in transcripts and predicted genes of diverse metabolic processes, particularly those involved in the catabolism of diverse organic substrates, including through fermentation, implicating these periods as important for microbial degradation of dietary compounds. These findings also suggest the period shortly after feeding may be a time of increased bacteria-bacteria and bacteria-host interaction. The movement of food into the intestine may stimulate bacteria living in association with the host mucus layer to mobilize and attach to food particles. These early responders therefore may be those cells best equipped to navigate a dynamic and spatially structured environment. Such drastic changes in the ecology of the microbiome around a

feeding cycle have not been demonstrated until now. Sampling the gut microbiome from other host species around a feeding cycle will help determine if these patterns are consistent and if microbes of these various lifestyles (food-induced bloomers vs. stable community members) are similar across diverse hosts.

5.3 Single cell genome sequencing to isolate gut taxa

In Appendix A, I begin exploring some of the taxa implicated in Ch. 3 and Ch. 4 as being important to the gut microbiome of marine fish. Isolating members of the gut microbiome for targeted study is difficult since most gut consitutuents cannot be cultured in the lab. Metagenomic and metatranscriptomic sequencing of the microbial fraction from gut communities is also hard because the pool of DNA/RNA from gut extracts is usually dominated (>95%) by host signal. Single cell amplified genome (SAG) sequencing may be the best approach for isolating members of these communities. The dataset in Appendix A represents the first SAG data from the gut of a marine fish. I used this dataset to begin exploring the diversity and function of an *Endozoicomonas* genotype associated with the butterflyfish *Chaetodon lunulatus*. This taxon has been described as a potential symbiont in marine invertebrate microbiomes but until now has not been described in fish.

Annotation and phylogenetic placement of the draft *Endozoicomonas* genome presented in Appendix A identified a metabolically flexible heterotroph with a functional repertoire similar to previously described Endozoicomonad genotypes. For a strict corallivorous fish like the putative host, secondary production of metabolites and vitamins by microbiome constituents might be especially important since corals are considered a nutritionally poor diet (Berumen et al. 2011). Given the high connectivity

between corals and coral feeding fish, it is also possible this isolate represents a coral-associated, transient member of the fish gut whose presence is merely a reflection of diet. However, the reverse case could also be made for many of the Endozoicomonads in corals. Indeed, the detection of Endozoicomonads in non-coral feeding fish (Chapter 3) and in other distant marine invertebrates would suggest that these organisms are adaptable to a wide range of hosts. The novel genotype sequenced and described here will serve as a valuable starting point for comparative studies of Endozoicomonads in other vertebrate gut microbiomes. Similarly, this SAG dataset as a whole serves as a starting point for uncovering the functional role of these microbes in fish systems and for comparison of fish-associated microbes to those isolated from other vertebrate hosts.

5.4 Final remarks

Fish gut microbiomes are diverse and highly variable as a result of the combined effects of many different organizing factors. However, there is likely a small number of commensal and symbiotic microbial taxa actually important to host health and ecology. The goal of this dissertation was to pare down the complexity of gut systems and identify key players in the microbiome of marine fish. This was accomplished by focusing on shifts associated with a major life stage transition in reef fish and examining short-term effects due to feeding on gut communities in clownfish. The trends identified here will serve as a valuable starting point for future, targeted studies on the potential biological and ecological significance of individual taxa in the gut.

APPENDIX A

DESCRIPTION OF A NOVEL ENDOZOICOMONAS SPECIES ISOLATED FROM INTESTINES OF THE OVAL BUTTERFLYFISH (*CHAETODON LUNULATUS*)

A.1 Introduction

Recognition of the role of microbiomes in animal health, behavior, and ecology has led to an increase in targeted genomic studies evaluating the functional potential of microbes associated with eukaryotic hosts (Baumgarten et al. 2015, Ding et al. 2016, Neuvonen et al. 2016). Gamma-proteobacteria of the Endozoicomonas genus were first described in 2007 following isolation of *Endozoicomonas elysicola* from the intestinal tract of the marine sea slug *Elysicola ornate* (Kurahashi and Yokota, 2007). Since this initial characterization, members of this genus have been shown to associate with diverse marine invertebrates including, sponges, corals, annelids, and bivalves across all the world's oceans (Neave et al. 2017A). They are particularly abundant in molecular surveys of the coral microbiome comprising up to 60% of recovered sequences and their abundance is positively correlated with coral health (Bayer et al. 2013, Bourne et al. 2013, Glayl et al. 2016). Recently, Endozoicomonads have also been described in vertebrate systems as core members of the gut microbiome in cardinalfish and damselfish and as causative agents of lesions in the sharpsnout seabream (Katharios et al. 2015, Parris et al. 2016). This genus is clearly emerging as a ubiquitous constituent of marine microbiomes, however, their diversity and function is unexplored for the majority of potential host taxa.

Our current knowledge on Endozoicomonads comes from several host-specific genotypes isolated or sequenced from marine corals, sponges, and most recently skin legions in seabream (Pike et al. 2013, Hyun et al. 2014, Appolinario et al. 2016, Schrieber et al. 2016, Katharios et al. 2015). Despite numerous investigations, there is no consensus on their primary functional role in microbiome systems. Endozoicomonad bacteria are most often described as symbionts and have been observed within tissues and cells of corals and bivalves where they are believed to participate in a nutritional symbiosis by transforming different organic carbon sources into substrates that can be used by the host (Neave et al. 2014, Ding et al. 2016, Neave et al. 2016, Neave et al. 2017B). Other functions ranging from sulfur cycling to protection from pathogens via antibiotics have been suggested, however, there is no direct evidence to support these hypotheses and conclusions are based primarily on the presence or absence of certain gene pathways within the putative symbiont genome (Neave et al. 2017A).

Much of the uncertainty over the function of Endozoicomonads arises from the fact their genomes are large (5.6-6.83 Mbs) and diverse, encoding metabolic pathways for both catabolism and biosynthesis of an array of carbohydrates, lipids, and amino acids (Pike et al. 2013, Hyun et al. 2014, Neave et al. 2014, Appolinario et al. 2016, Neave et al. 2017A). These genomes support a similarly large number of protein-coding genes suggesting that genome streamlining which occurs in obligate symbionts is not happening in these taxa. Another feature common to all sequenced Endozoicomonad genomes is the presence of a high number of transposable elements and repeat sequences which may help them adapt from a free-living state to niches within diverse eukaryotic hosts (Qi et al. 2018). Finally, Endozoicomonad genomes encode many genes in the

phosphoenolpyruvate (PEP)-dependent phosphotransferase system (PTS) used for uptake of carbohydrate substrates and secretory pathways (ABC transporters and secretion systems) for potentially translocating transformed substrates back outside of the cell for host use (Neave et al. 2014). While this genomic evidence hints at a functional role for Endozoicomonads in microbiomes, their prevalence and diversity in other host systems is still largely unknown. Characterization of Endozoicomads outside of marine invertebrates may help provide additional insight on the potential functional role and ubiquity of these taxa.

The coral reef habitats where Endozoicomonads have been detected in invertebrates harbor over 2500 fish species engaged in a complex web of interactions, including predation, herbivory, corallivory, and symbiosis (Bellwood et al. 2004; Allen 2014). The oval butterflyfish, *Chaetodon lunulatus*, is a coral feeding specialist whose diet is comprised almost exclusively of hard coral (Pratchett, 2005). As a result of this feeding strategy, their gut is likely inundated with high densities of Endozoicomonad bacteria. We hypothesized fish gut microbiomes might harbor distinct genotypes of these bacteria specifically adapted for this type of microbiome. Rather than true symbionts, perhaps Endozoicomads are generalist heterotrophs suited for life in the organic rich, biofilm communities provided by microbiomes with their role being more commensal than symbiotic. The data that follow describe a nearly complete, high quality Endozoicomonas genome isolated from the gut of the oval butterflyfish via single cell sequencing. We show that this bacterium is phylogenetically distinct from previously described genotypes with a similarly diverse functional repertoire. These results identify

a potentially fish specific Endozoicomonad species and serves as a starting point for comparative studies of other vertebrate-associated Endozoicomad genotypes.

A.2 Methods and Materials

Oval butterfly fish (*Chaetodon lunulatus)* were field-collected from the island of Moorea in French Polynesia in February 2016 via hand net as part of a separate study (Pratte et al. 2018). Specimens were immediately euthanized using a lethal dose of tricaine mesylate (MS-222) and transported back to a field lab on ice. In the lab, oval butterflyfish were dissected aseptically and whole intestines removed. A longitudinal cut was made in the intestine and a slurry prepared by washing 1-2 mL of sterile water over the intestinal contents and collecting in a sterile tube. Due to the viscosity of this slurry, we further prepared a 1 in 100 dilution of sample. Diluted sample was preserved by adding 1000 µL to 100 µL glycerol-TE cryoprotectant stock (20 mL 100x TE pH 8.0, 60 mL DI water, 100 mL molecular-grade glycerol). Samples were stored in a freezer at -20 degrees Celsius. A single sample was sent to Bigelow Single Cell Genomics Center for cell sorting and generation of single cell amplified genomes (Yilmaz and Singh, 2012).

Single cell amplified genome (SAG) generation yielded 78 samples positive for 16S rRNA (i.e. non-host). These samples represent the first SAGs isolated from the gut of marine fish. Of these SAGs, 2 were classified to the Endozoicomonas genus with >90% certainty via RDP and their 16S sequences were identical. Barcoded Illumina sequence libraries were prepared from these samples using the Nextera kit and sequenced on 2 separate 250bp x 250bp Miseq runs. Sequencing yielded approximately 13 GB of data. Reads were screened for sequencing adapter and quality trimmed using Trimgalore (Krueger et al. 2015) to remove any bases below phred score 30 and sequences shorter

than 150bp. Unmerged reads were assembled in SPAdes 3.13 using single-cell mode and the draft assembly generated was used for downstream analysis (Bankevich et al. 2012). Genome completeness and contamination based on single-copy genes was estimated using QUAST (Gurevich et al. 2013). Genes were predicted from the genome assembly using Prodigal (Hyatt et al. 2010).

Phylogeny of the butterflyfish isolate was examined through comparison to 12 complete or nearly complete Endozoicomonas assemblies in the NCBI database (Table A.1). 139 single copy marker genes identified in Campbell et al. (2011) and conserved across all genomes were aligned with MUSCLE (Edgar, 2004) and concatenated in Anvio (Eren et al. 2015). Concatenated protein alignments were exported into MEGA7 (Kumer et al. 2016) and phylogeny modeled using the Maximum likelihood method with bootstrap values calculated based on 100 iterations. Amino acid Identity (AAI) was calculated for my genome vs. every available Endozoicomonas assembly via AAI calculator in the enveomics toolbox (Rodriguez-R and Konstantinidis 2016). Predicted proteins were annotated via blast and ghostx with Blastkoala (Kanehisa et al. 2016).

Table A.1. Accession numbers and assembly size for previously published Endozoicomonad genome assemblies

Name	Strain	Accession Number	Size (Mbp)
Endozoicomonas elysicola	DSM 22380	GCF_000710775.1	5.61
Endozoicomonas atrinae	WP70	GCF_001647025.1	6.69
Endozoicomonas montiporae	CL-33(T)	GCF_000722565.1	5.43
Endozoicomonas arenosclerae	E_MC227	GCA_001562005.1	6.22
Endozoicomonas numazuensis	DSM 25634	GCF_000722635.1	6.34
Endozoicomonas montiporae	LMG 24815	GCF_000722565.1	5.6
Endozoicomonas arenosclerae	Ab112	GCF_001562015.1	6.45
Endozoicomonas cretensis	DpTailN28	GCA_900299555.1	5.88
Endozoicomonas ascidiicola	AVMART05	GCF_001646945.1	6.13
Endozoicomonas ascidiicola	KASP37	GCF_001646955.1	6.51
Endozoicomonas sp.	AB1-5	GCA_001729985.1	4.049
Endozoicomonas acroporae	ACR-14	GCA_002864045.1	6.05

A.3 Results and Discussion

I assembled a 3,315,008 base pair Endozoicomonas genome from single cells isolated from the intestine of the oval butterflyfish, *Chaetodon lunulatus*. RDP classification placed this draft genome with 93% confidence into the Endozoicomonas genus. Genome completeness was estimated at 80.5 % with low contamination and the GC content of 53.12% was within the range observed for other Endozoicomonad genotypes (46%-54%, Table A.2). Based on the completion estimate, this genome is smaller than the genomes of previously described Endozoicomonas genotypes which range from 5.6-6.82 Mbs (Neave et al. 2014).

Table A.2 Spades assembly statistics for Endozoicomonas genome.

Contigs	N50	Length	GC content	Predicted proteins	Coding density	Completeness	Contamination
510	14248 bp	3.315 Mbp	53.12%	3,095	86%	80.50%	7.50%

To place this assembly into a phylogenetic context, I constructed an amino acid tree based on 139 single copy genes present in this isolate and all publicly available Endozoicomonad genomes (Figure A.1, Table A.1, Campbell et al. 2011). There was some support for host phylogeny reflecting Endozoicomonad phylogeny as 5 of 6 coral and sponge associated genotypes clustered together separately (Figure A.1). However, the Endozoicomonas genotype in this study was part of a monophyletic clade of isolates from diverse hosts including a sea slug, bryozoan, coral, bivalve, and fish (Figure A.1). The genotypes within this clade shared approximately 70% AAI with my draft genome indicating that this sequence likely represents a novel species (Table A.3, Konstantinidis and Tiedje, 2005). To date, no more than 78% AAI has been observed between any two

Endozoicomonad genotypes highlighting the high genomic diversity within this genus (Appolinario et al. 2016).

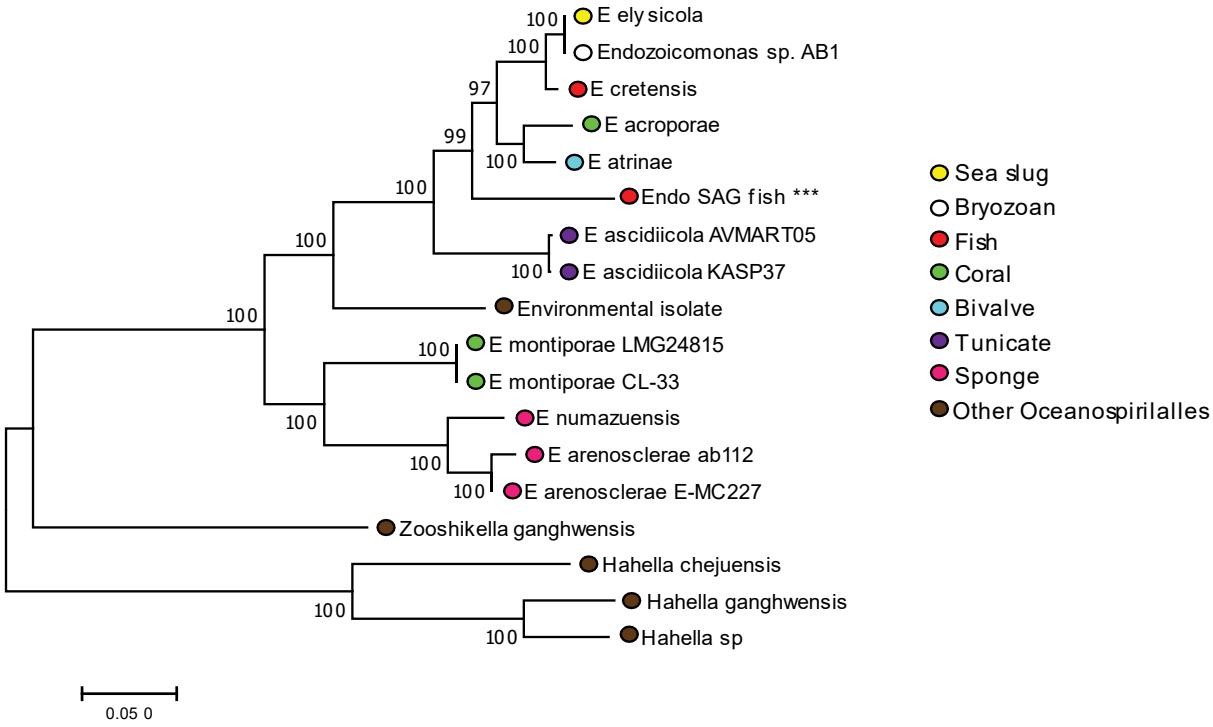

Figure A.1 Phylogeny of Endozoicomonad isolates based on 139 single-copy genes conserved in all draft genomes. Evolutionary history was inferred from amino acid sequences using the Maximum likelihood method based on the JTT matrix model and bootstrap values are based on 100 iterations. There were 15,276 positions in the final dataset. Colored circles represent host type with non-hosting members of the Oceanospirillales included as an outgroup. The Endozoicomonas draft genome in this study is labeled "Endo SAG fish" and denoted with stars. All evolutionary analyses were conducted in MEGA7 (Kumar et al. 2016).

Table A.3 Average amino acid identity (AAI) suggests the Endozoicomonas genotype isolated from the gut of the oval butterfly fish represents a new species. AAI was estimated using average amino acid identity from best hits (one-way AAI) and reciprocal best hits (two-way AAI) between two genomic datasets of proteins (Rodriguez-R and Konstantinidis, 2014).

Genome A	Genome B	AAI	SD	N*	Omega**	Frx***
Endo_SAG_fish	E. acroporae	70.01441	19.94855	1887	3095	60.96931
Endo_SAG_fish	E. arenosclerae E-MC227	55.06402	19.40232	1711	3095	55.28271
Endo_SAG_fish	E. arenosclerae ab112	58.89947	18.76736	1797	3095	58.06139
Endo_SAG_fish	Endozoicomonas sp. AB1	70.6235	19.86818	1828	3095	59.063
Endo_SAG_fish	E. ascidiicola AVMART05	65.9357	20.04538	1899	3095	61.35703
Endo_SAG_fish	E. ascidiicola KASP37	66.25548	19.79015	1904	3095	61.51858
Endo_SAG_fish	E. atrinae	70.08693	20.4994	1904	3095	61.51858
Endo_SAG_fish	E. cretensis	70.12811	19.88704	1871	3095	60.45234
Endo_SAG_fish	E. elysicola	70.51658	19.9388	1833	3095	59.22456
Endo_SAG_fish	E. montiporae CL-33	59.55789	18.34141	1779	3095	57.47981
Endo_SAG_fish	E. montiporae LMG 24815	59.64322	18.35149	1765	3095	57.02746
Endo_SAG_fish	E. numazuensis	59.63211	18.33196	1743	3095	56.03247

* Number of reciprocal best matches found

**Minimum # of proteins between the two genomes

***Percentage of genome shared

To characterize the functional potential of my assembly, proteins were predicted using Prodigal (Hyatt et al. 2010). Gene prediction estimated 3,095 coding sequences in the draft genome (Table A.2). Of those proteins, about half (1,477) could be annotated into KEGG. KEGG pathways containing the most gene hits included those associated with carbohydrate and amino acid metabolism, membrane transport systems associated with the movement of carbohydrates and amino acids, and biosynthesis of cofactors and vitamins (Figure A.2). Approximately 20% of predicted proteins in the genome were dedicated to these pathways alone. Components of the PTS system (used for uptake of carbohydrates) included genes for the acquisition of glucose, sucrose, and cellobiose. Cellobiose is a major component of algal cells and it has been posited that enrichment of genes for cellobiose metabolism in Endozoicomonads (compared to other

Oceanospirillales) may indicate a niche for recycling carbon in decaying *Symbiodinium* cells in corals (Neave et al. 2017A). There were also many components of the Type II, Type III, and Sec-SRP secretions systems involved in translocation of proteins from inside to outside of the cell and in potential interactions with host cells (Ding et al. 2016). The bulk trends in metabolic capability discussed above have been demonstrated for Endozoicomonad genotypes isolated from marine invertebrates (Neave et al. 2016) and suggest these diverse taxa are functionally similar across host systems.

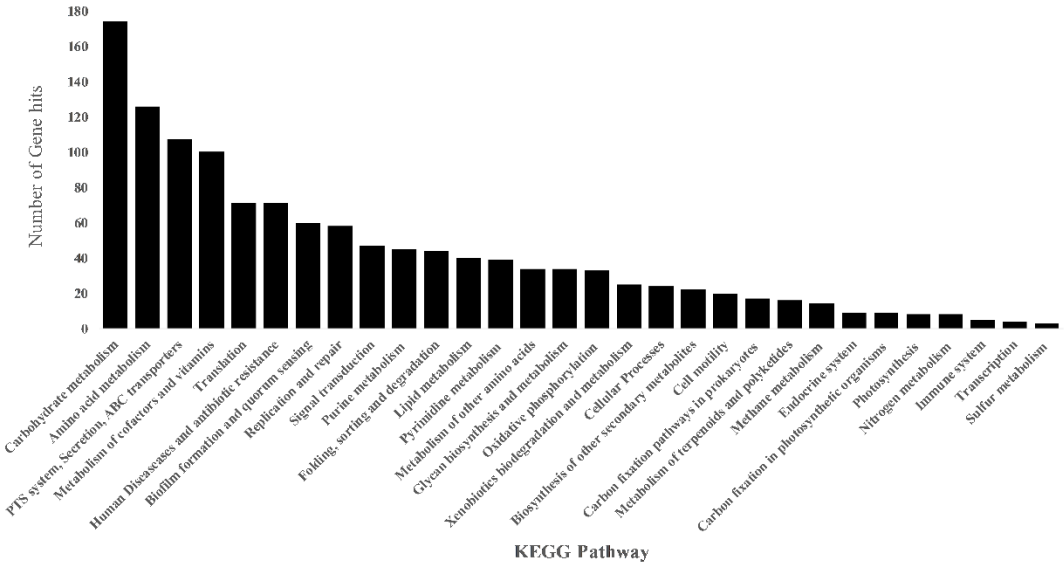

Figure A.2 The highest proportion of genes in the Endozoicomonas draft genome are associated with carbohydrate metabolism, amino acid metabolism, and protein transport. Shown are the number of gene hits in the genome assigned to large KEGG categories. Annotations were made via blast and ghostx searches against the KEGG database using BlastKoala (Kanehisa et al. 2016). Nearly half of all predicted proteins were annotated.

Several interesting non-metabolic features were also observed in the Endozoicomonas genome described here including 67 genes involved in quorum sensing and biofilm formation. These systems may help Endozoicomonads organize themselves

into the cyst-like growth form they exhibit in some marine invertebrates and during pathogenesis in fish (Medoza et al. 2013, Neave et al. 2017A). The mechanistic basis for formation and maintenance of these cysts is still unknown. Also present in this Endozoicomonas genome were 29 transposases and the full suite of genes for homologous recombination in bacteria. Given the large size of Endozoicomonad genomes, it is likely that they have a free-living state and may even transition between free-living and host-associated lifestyles regularly. The high frequency of mobile elements and recombination genes observed here and in previously described Endozoicomonas taxa may help them adapt quickly to diverse host environments.

A.4 Conclusions

Annotation and phylogenetic placement of the draft genome presented here identifies a metabolically flexible heterotroph with a functional repertoire similar to previously described Endozoicomonad genotypes. For a strict corallivore like the putative host, *Chaetodon lunulatus*, secondary production of metabolites and vitamins by microbiome constituents might be especially important since corals are considered a nutritionally poor diet (Berumen et al. 2011). Given the high connectivity between corals and coral feeding fish, it is also possible this isolate represents a coral-associated, transient member of the fish gut whose presence is merely a reflection of diet. However, the reverse case could also be made for many of the Endozoicomonads in corals. Indeed, the detection of Endozoicomonads in non-coral feeding fish (Katharios et al. 2015, Parris et al. 2016) and in other distant marine invertebrates would suggest that these organisms are adaptable to a wide range of hosts. The novel genotype sequenced and described here

will serve as a valuable starting point for comparative studies of Endozoicomonads in other vertebrate gut microbiomes.

A.5 References

Allen, G. R. (2014). Review of Indo-Pacific coral reef fish systematics: 1980 to 2014. *Ichthyological Research*, 62(1), pp. 2-8.

Appolinario, L. R., Tschoeke, D. A., Rua, C. P., Venas, T., Campeão, M. E., Amaral, G. R., ... & Swings, J. (2016). Description of Endozoicomonas arenosclerae sp. nov. using a genomic taxonomy approach. *Antonie van Leeuwenhoek*, *109*(3), 431-438.

Baumgarten, S., Simakov, O., Esherick, L. Y., Liew, Y. J., Lehnert, E. M., Michell, C. T., ... & Gough, J. (2015). The genome of Aiptasia, a sea anemone model for coral symbiosis. *Proceedings of the National Academy of Sciences*, 112(38), 11893-11898.

Bayer, T., Arif, C., Ferrier-Pagès, C., Zoccola, D., Aranda, M., & Voolstra, C. R. (2013). Bacteria of the genus Endozoicomonas dominate the microbiome of the Mediterranean gorgonian coral Eunicella cavolini. Marine Ecology Progress Series, 479, 75-84.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., ... & Pyshkin, A. V. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology*, *19*(5), 455-477.

Bellwood DR, Hughes TP, Folke C, & Nyström M. (2004). Confronting the coral reef crisis. *Nature*. 429, pp. 827- 833.

Berumen, M. L., Pratchett, M. S., & Goodman, B. A. (2011). Relative gut lengths of coral reef butterflyfishes (Pisces: Chaetodontidae). *Coral Reefs*, 30(4), 1005.

Bourne, D. G., Dennis, P. G., Uthicke, S., Soo, R. M., Tyson, G. W., & Webster, N. (2013). Coral reef invertebrate microbiomes correlate with the presence of photosymbionts. The ISME journal, 7(7), 1452.

Campbell, B. J., Yu, L., Heidelberg, J. F., & Kirchman, D. L. (2011). Activity of abundant and rare bacteria in a coastal ocean. *Proceedings of the National Academy of Sciences*, *108*(31), 12776-12781.

Ding, J. Y., Shiu, J. H., Chen, W. M., Chiang, Y. R., & Tang, S. L. (2016). Genomic insight into the host–endosymbiont relationship of Endozoicomonas montiporae CL-33T with its coral host. *Frontiers in microbiology*, 7, 251.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, *32*(5), 1792-1797.

Eren, A. M., Esen, Ö. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., & Delmont, T. O. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ*, *3*, e1319.

Glasl, B., Herndl, G. J., & Frade, P. R. (2016). The microbiome of coral surface mucus has a key role in mediating holobiont health and survival upon disturbance. The ISME journal, 10(9), 2280.

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, *29*(8), 1072-1075.

Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics*, *11*(1), 119.

Hyun, D. W., Shin, N. R., Kim, M. S., Oh, S. J., Kim, P. S., Whon, T. W., & Bae, J. W. (2014). Endozoicomonas atrinae sp. nov., isolated from the intestine of a comb pen shell Atrina pectinata. *International journal of systematic and evolutionary microbiology*, *64*(7), 2312-2318.

Kanehisa, M., Sato, Y., & Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of molecular biology*, 428(4), 726-731.

Katharios, P., Seth-Smith, H. M., Fehr, A., Mateos, J. M., Qi, W., Richter, D., ... & Thomson, N. R. (2015). Environmental marine pathogen isolation using mesocosm culture of sharpsnout seabream: striking genomic and morphological features of novel Endozoicomonas sp. *Scientific reports*, 5, 17609.

Konstantinidis, K. T., & Tiedje, J. M. (2005). Towards a genome-based taxonomy for prokaryotes. *Journal of bacteriology*, *187*(18), 6258-6264.

Krueger, F. (2015). Trim galore. *A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files*.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular biology and evolution*, *33*(7), 1870-1874.

Kurahashi, M., & Yokota, A. (2007). Endozoicomonas elysicola gen. nov., sp. nov., a γ-proteobacterium isolated from the sea slug Elysia ornata. *Systematic and applied microbiology*, 30(3), 202-206.

Mendoza, M., Güiza, L., Martinez, X., Caraballo, X., Rojas, J., Aranguren, L. F., & Salazar, M. (2013). A novel agent (Endozoicomonas elysicola) responsible for epitheliocystis in cobia Rachycentrum canadum larvae. *Diseases of aquatic organisms*, *106*(1), 31-37.

Neave, M. J., Michell, C. T., Apprill, A., & Voolstra, C. R. (2014). Whole-genome sequences of three symbiotic Endozoicomonas strains. *Genome announcements*, *2*(4), e00802-14.

Neave, M. J., Apprill, A., Ferrier-Pagès, C., & Voolstra, C. R. (2016). Diversity and function of prevalent symbiotic marine bacteria in the genus Endozoicomonas. *Applied microbiology and biotechnology*, 100(19), 8315-8324.

Neave, M. J., Michell, C. T., Apprill, A., & Voolstra, C. R. (2017A). Endozoicomonas genomes reveal functional adaptation and plasticity in bacterial strains symbiotically associated with diverse marine hosts. *Scientific reports*, *7*, 40579.

Neave, M. J., Rachmawati, R., Xun, L., Michell, C. T., Bourne, D. G., Apprill, A., & Voolstra, C. R. (2017B). Differential specificity between closely related corals and abundant Endozoicomonas endosymbionts across global scales. *The ISME journal*, *11*(1), 186.

Neuvonen, M. M., Tamarit, D., Näslund, K., Liebig, J., Feldhaar, H., Moran, N. A., ... & Andersson, S. G. (2016). The genome of Rhizobiales bacteria in predatory ants reveals urease gene functions but no genes for nitrogen fixation. *Scientific reports*, 6, 39197.

Parris, D. J., Brooker, R. M., Morgan, M. A., Dixson, D. L., & Stewart, F. J. (2016). Whole gut microbiome composition of damselfish and cardinalfish before and after reef settlement. PeerJ, 4, e2412.

Pike, R. E., Haltli, B., & Kerr, R. G. (2013). Description of Endozoicomonaseuniceicola sp. nov. and Endozoicomonas gorgoniicola sp. nov., bacteria isolated from the octocorals Eunicea fusca and Plexaura sp., and an emended description of the genus Endozoicomonas. *International journal of systematic and evolutionary microbiology*, *63*(11), 4294-4302.

Pratchett, M. S. (2005). Dietary overlap among coral-feeding butterflyfishes (Chaetodontidae) at Lizard Island, northern Great Barrier Reef. *Marine Biology*, 148(2), 373-382.

Pratte, Z. A., Besson, M., Hollman, R. D., & Stewart, F. J. (2018). The gills of reef fish support a distinct microbiome influenced by host-specific factors. *Applied and environmental microbiology*, AEM-00063.

Qi, W., Cascarano, M. C., Schlapbach, R., Katharios, P., Vaughan, L., & Seth-Smith, H. M. (2018). Ca. Endozoicomonas cretensis: A Novel Fish Pathogen Characterized by Genome Plasticity. *Genome biology and evolution*, *10*(6), 1363-1374.

Rodriguez-R, L. M., & Konstantinidis, K. T. (2014). Bypassing cultivation to identify bacterial species. *Microbe*, *9*(3), 111-118.

Rodriguez-R, L. M., & Konstantinidis, K. T. (2016). *The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes* (No. e1900v1). PeerJ Preprints.

Schreiber, L., Kjeldsen, K. U., Obst, M., Funch, P., & Schramm, A. (2016). Description of Endozoicomonas ascidiicola sp. nov., isolated from Scandinavian ascidians. *Systematic and applied microbiology*, *39*(5), 313-318.

Yilmaz, S., & Singh, A. K. (2012). Single cell genome sequencing. *Current opinion in biotechnology*, *23*(3), 437-443.

Appendix B

Supplementary information for Chapter 3: Whole gut microbiome composition of damselfish and cardinalfish before and after reef settlement

Table B.1 Summary of fish samples collected and analyzed

Family	Genus	Species	Life Stage	No. of sequences	OTUs post rarefaction	Fish Length (mm)
Pomacentridae	Pomacentrus	moluccensis	Pre-settlement	9848	102	13
Pomacentridae	Pomacentrus	moluccensis	Pre-settlement	4932	155	12
Pomacentridae	Pomacentrus	moluccensis	Pre-settlement	7610	557	11
Pomacentridae	Pomacentrus	moluccensis	Pre-settlement	9197	250	11
Pomacentridae	Pomacentrus	chrysurus	Pre-settlement	64390	148	15
Pomacentridae	Pomacentrus	chrysurus	Pre-settlement	19825	135	14
Pomacentridae	Pomacentrus	chrysurus	Pre-settlement	22546	107	14
Pomacentridae	Pomacentrus	chrysurus	Pre-settlement	10228	170	14
Pomacentridae	Pomacentrus	nagasakiensis	Pre-settlement	20696	61	16
Pomacentridae	Pomacentrus	nagasakiensis	Pre-settlement	28458	107	16
Pomacentridae	Pomacentrus	nagasakiensis	Pre-settlement	11828	110	15
Pomacentridae	Pomacentrus	nagasakiensis	Pre-settlement	34479	69	16
Pomacentridae	Pomacentrus	amboinensis	Pre-settlement	5887	216	14
Pomacentridae	Pomacentrus	amboinensis	Pre-settlement	10060	253	14
Pomacentridae	Pomacentrus	amboinensis	Pre-settlement	74686	131	13
Pomacentridae	Pomacentrus	amboinensis	Pre-settlement	25807	303	16
Apogonidae	unknown	unknown	Pre-settlement	13096	158	8
Apogonidae	unknown	unknown	Pre-settlement	80429	106	8
Apogonidae	unknown	unknown	Pre-settlement	10393	139	8
Apogonidae	unknown	unknown	Pre-settlement	31293	164	9
Apogonidae	unknown	unknown	Pre-settlement	36882	375	8
Apogonidae	unknown	unknown	Pre-settlement	22128	212	10
Apogonidae	unknown	unknown	Pre-settlement	30059	226	10
Pomacentridae	Pomacentrus	wardii	Pre-settlement	38610	251	14
Pomacentridae	Pomacentrus	wardii	Pre-settlement	82997	258	15
Pomacentridae	Pomacentrus	wardii	Pre-settlement	90795	141	15
Pomacentridae	Pomacentrus	wardii	Pre-settlement	138124	311	16
Pomacentridae	Pomacentrus	bankanensis	Pre-settlement	12279	63	13
Pomacentridae	Pomacentrus	bankanensis	Pre-settlement	8732	145	13
Pomacentridae	Pomacentrus	bankanensis	Pre-settlement	12333	164	14
Pomacentridae	Pomacentrus	bankanensis	Pre-settlement	7197	104	11
Pomacentridae	Pomacentrus	coelestis	Pre-settlement	22552	267	15
Pomacentridae	Pomacentrus	coelestis	Pre-settlement	7089	257	15
Pomacentridae	Pomacentrus	coelestis	Pre-settlement	8211	124	15
Pomacentridae	Pomacentrus	coelestis	Pre-settlement	25037	244	15
Pomacentridae	Dascyllus	aruanus	Pre-settlement	4398	178	9
Pomacentridae	Dascyllus	aruanus	Pre-settlement	16510	278	9
Pomacentridae	Dascyllus	aruanus	Pre-settlement	6478	180	9
Pomacentridae	Chromis	unknown	Pre-settlement	4440	171	11
Pomacentridae	Chromis	unknown	Pre-settlement	5366	193	12
Pomacentridae	Chromis	unknown	Pre-settlement	8756	178	11
Pomacentridae	Chromis	unknown	Pre-settlement	12667	187	11
Pomacentridae	Chromis	unknown	Pre-settlement	10637	99	11
Pomacentridae	Chromis	unkrown	Pre-settlement	42028	148	11
Pomacentridae	Chromis	unknown	Pre-settlement	8226	138	11
Apogonidae	Ostorhinchus	doederleini	Pre-settlement	59740	155	12
Apogonidae	Ostorhinchus	doederleini	Pre-settlement	13338	55	11
Apogonidae	Ostorhinchus	doederleini	Pre-settlement	5921	206	12
Apogonidae	Ostorhinchus	doederleini	Pre-settlement	6887	173	12
Pomacentridae	Dascyllus	aruanus	Post-settlement	31404	163	30
Pomacentridae	Dascyllus	aruanus	Post-settlement	4262	258	24
Apogonidae	Ostorhinchus	doederleini	Post-settlement	33602	297	29
Apogonidae	Ostorhinchus	doederleini	Post-settlement	16642	162	28
Apogonidae	Ostorhinchus	doederleini	Post-settlement	10612	287	32
Apogonidae	Ostorhinchus	doederleini	Post-settlement	56229	294	33
Pomacentridae	Pomacentrus	moluccensis	Post-settlement	29632	320	36
Pomacentridae	Pomacentrus	moluccensis	Post-settlement	71204	93	31
Pomacentridae	Pomacentrus	moluccensis	Post-settlement	8672	181	29
Pomacentridae	Pomacentrus	moluccensis	Post-settlement	26880	187	34
Pomacentridae	Pomacentrus	chrysurus	Post-settlement	18609	44	25
Pomacentridae	Pomacentrus	chrysurus	Post-settlement	30552	183	27
Pomacentridae	Pomacentrus	chrysurus	Post-settlement	54271	49	27
Pomacentridae	Acanthochromis	polyacanthus	Post-settlement	16725	422	14
Pomacentridae	Acanthochromis	polyacanthus	Post-settlement	5293	325	16
Pomacentridae	Acanthochromis	polyacanthus	Post-settlement	10162	351	15
Pomacentridae	Pomacentrus	moluccensis	Post-settlement	15679	228	24
Pomacentridae	Pomacentrus	moluccensis	Post-settlement	11487	124	26
Pomacentridae	Chromis	unknown	Post-settlement	18184	299	31
Pomacentridae	Chromis	unknown	Post-settlement	30348	318	25
Pomacentridae	Chromis	unknown	Post-settlement	25948	301	22
Pomacentridae	Dascyllus	aruanus	Post-settlement	27731	322	25
Pomacentridae	Dascyllus	aruanus	Post-settlement	18668	281	27
Pomacentridae	Dascyllus	aruanus	Post-settlement	45540	380	30!

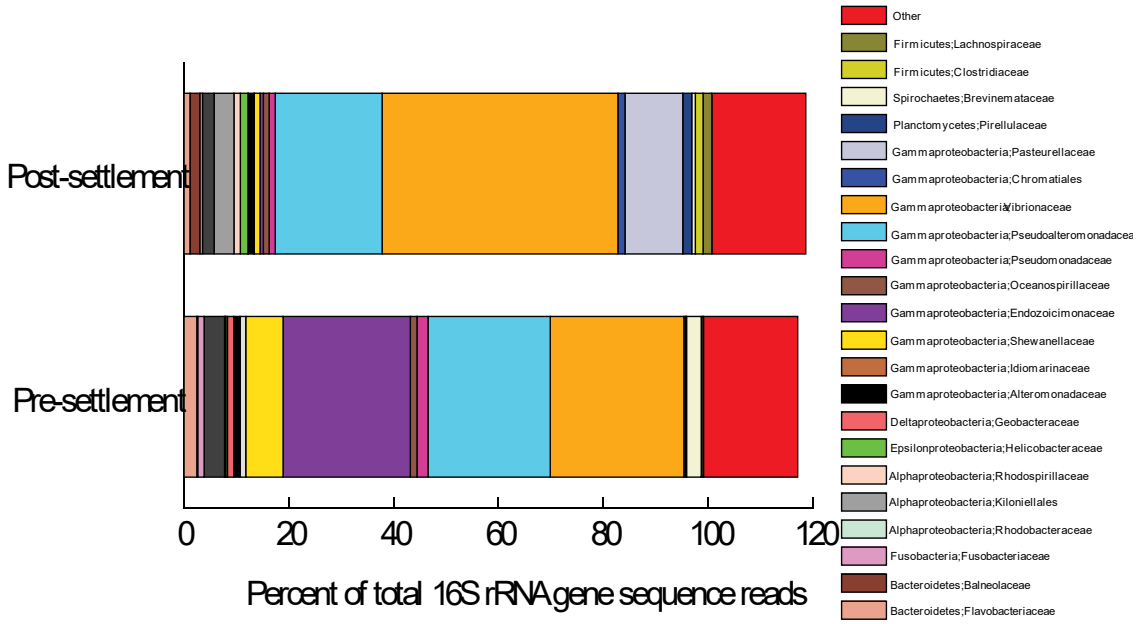

Figure B.1. Microbial community composition at the family level averaged across pre- (n=49) and post-settlement (n=24) damselfish and cardinalfish based on data rarified to 4262 sequences per sample.

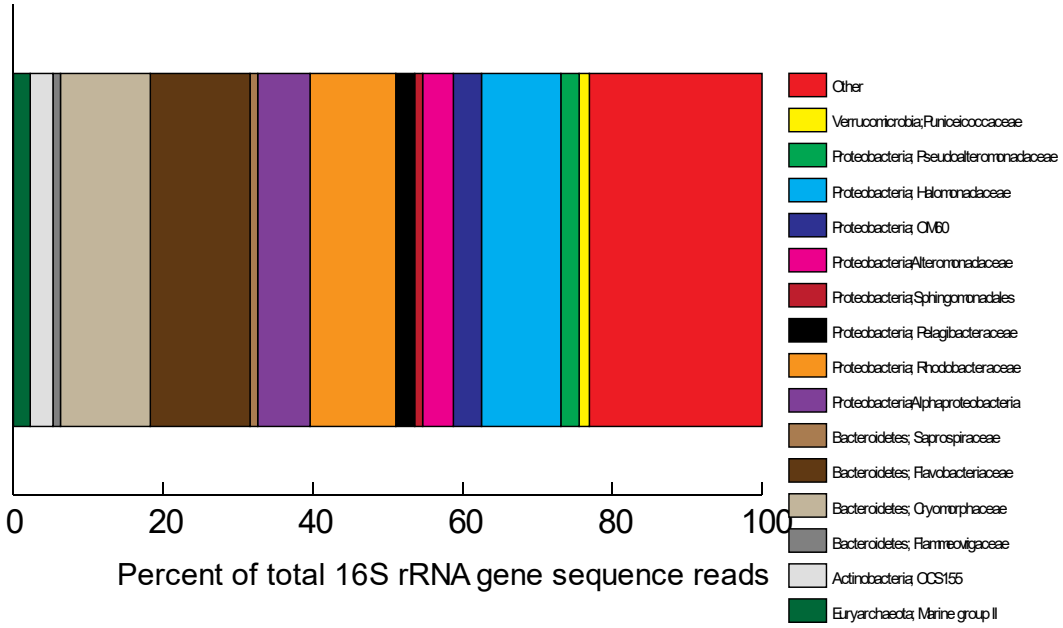

Figure B.2. Microbial community composition at the family level in seawater samples (n=11) with data rarified to 4262 sequences per sample.

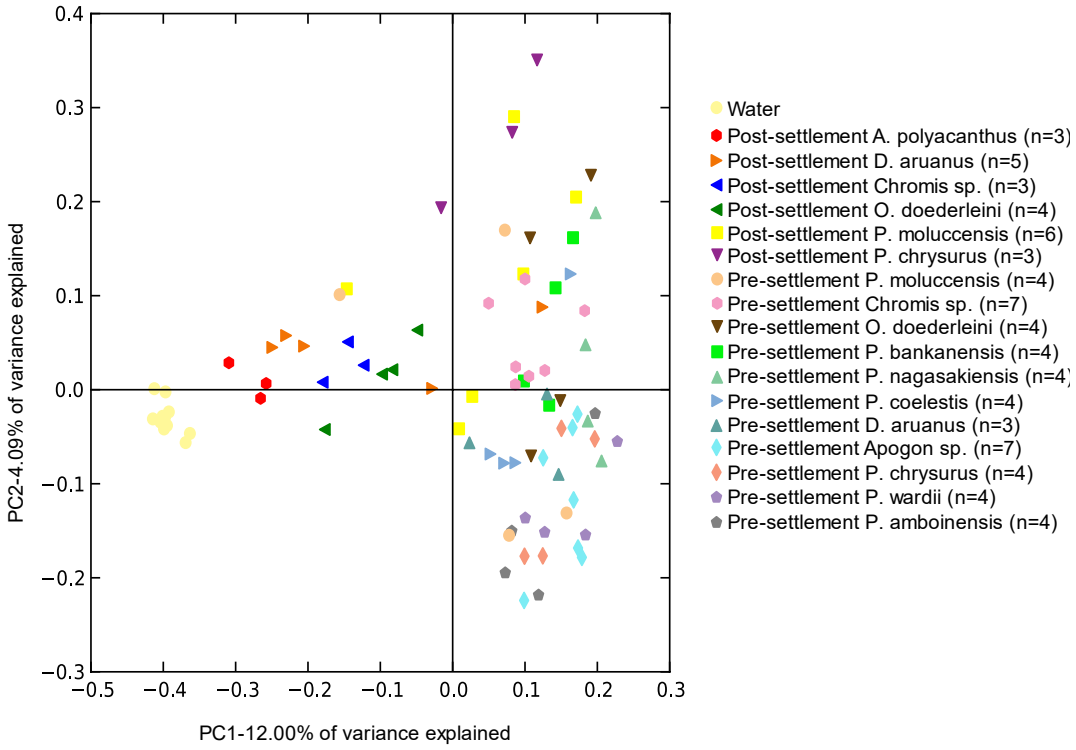

Figure B.3. **Compositional relatedness of microbiome samples**. Principal coordinate analysis based on unweighted Unifrac distance with sequence data rarified to 4262 sequences per sample (49 pre-settlement, 24 post-settlement, 11 seawater samples). Samples color-coded according to host.

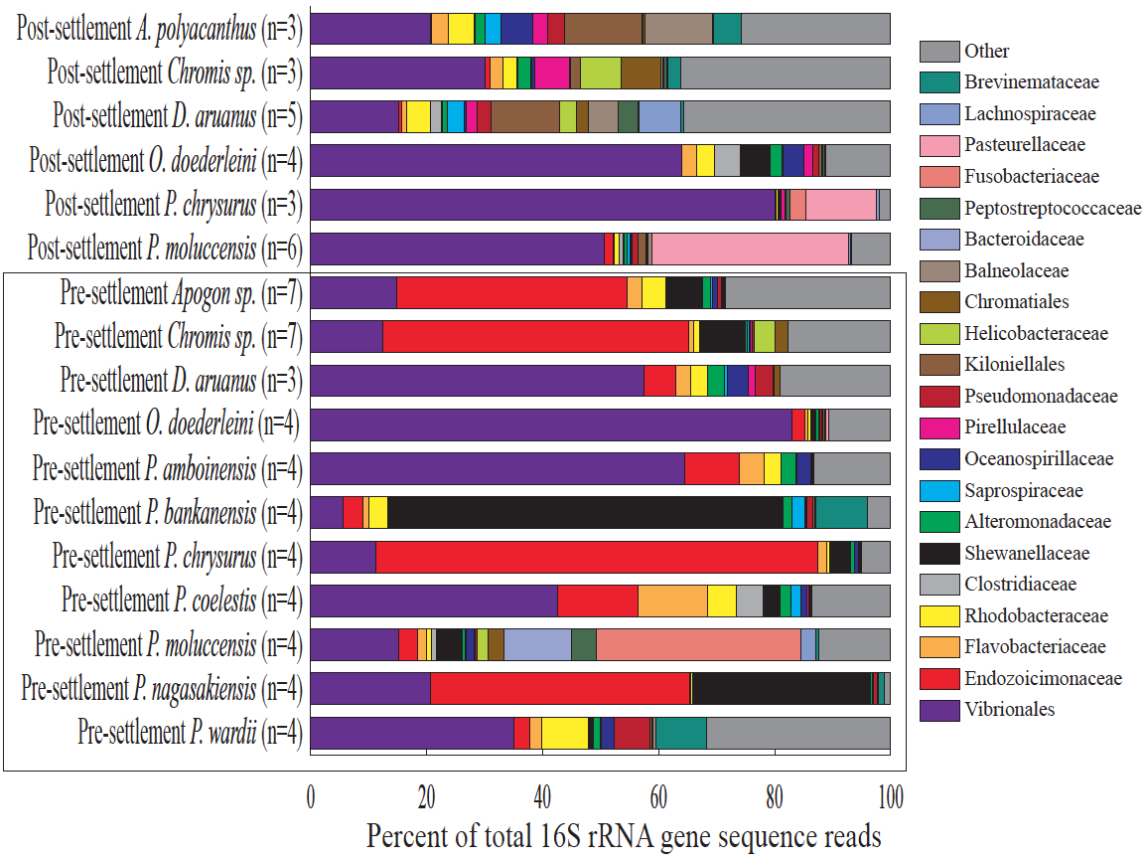

Figure B.4. **Microbial community composition based on data rarified to 4262 sequences per sample organized by host type**. Outlined host group represents pre-settlement individuals.

APPENDIX C

SUPPLEMENTARY INFORMATION FOR CHAPTER 4: FEEDING RAPIDLY ALTERS MICROBIOME COMPOSITION AND GENE TRANSCRIPTION IN THE CLOWNFISH GUT

Table C.1. Sequencing statistics and sample information for amplicon data following the Deblur method in Qiime2.

Sample #	Sample type	Treatment	Time	Raw # of reads	Deblur reads	Unique reads
1	Fish-experimental	4x	1230	144690	90649	408
2	Fish-experimental	4x	1230	111081	73221	386
3	Fish-experimental	4x	1230	149575	90093	653
4	Fish-experimental	1x	1400	128606	87375	278
5	Fish-experimental	1x	1400	127423	85954	551
6	Fish-experimental	1x	1400	82560	51285	312
7	Fish-experimental	0.33x	1400	111323	76632	333
8	Fish-experimental	0.33x	1400	136366	85919	357
9	Fish-experimental	0.33x	1400	80283	53619	285
10	Fish-experimental	4x	1400	34238	22054	260
11	Fish-experimental	4x	1400	104698	69542	470
12	Fish-experimental	4x	1400	113264	71535	650
13	Fish-experimental	1x	1600	108641	71170	517
14	Fish-experimental	1x	1600	93840	66429	408
15	Fish-experimental	1x	1600	36677	21867	254
16	Fish-experimental	0.33x	1600	89712	55871	446
17	Fish-experimental	0.33x	1600	78938	47337	313
18	Fish-experimental	0.33x	1600	103320	71406	169
19	Fish-experimental	4x	1600	66287	44214	342
20	Fish-experimental	4x	1600	53693	35660	155
21	Fish-experimental	4x	1600	38315	23484	183
22	Fish-experimental	1x	2000	75411	55138	76
23	Fish-experimental	1x	2000	107915	76200	235
24	Fish-experimental	1x	2000	55470	36872	267
25	Fish-experimental	0.33x	2000	112899	82765	179
26	Fish-experimental	0.33x	2000	100008	69795	233
27	Fish-experimental	0.33x	2000	70216	48742	102
28	Fish-experimental	4x	2000	53630	33315	273
29	Fish-experimental	4x	2000	15817	10354	141
30	Fish-experimental	4x	2000	81554	57491	197
31	Fish-experimental	1x	1230	93769	61198	445
32	Fish-experimental	1x	1230	61780	35839	209
33	Fish-experimental	1x	1230	114386	73895	418
34	Fish-experimental	0.33x	1230	124852	80507	607
35	Fish-experimental	0.33x	1230	128899	80254	388
36	Fish-experimental	0.33x	1230	126447	80521	360
37	Fish-experimental	4x	1100	47664	30747	181
38	Fish-experimental	4x	1100	51118	36966	96
39	Fish-experimental	4x	1100	81567	58231	276
40	Fish-experimental	1x	1100	56093	41300	124
41	Fish-experimental	1x	1100	82064	48668	169
42	Fish-experimental	1x	1100	1293	706	37
43	Fish-experimental	0.33x	1100	58712	43316	302
44	Fish-experimental	0.33x	1100	60079	40364	141
45	Fish-experimental	0.33x	1100	79423	46646	123
46	Fish-experimental	4x	2000	80677	55790	101
47	Fish-experimental	4x	2000	116208	82917	206
48	Fish-experimental	4x	2000	112020	80198	374
49	Fish-experimental	0.33x	2000	79260	59861	192
50	Fish-experimental	0.33x	2000	75799	58076	97
51	Fish-experimental	0.33x	2000	50038	38682	223
52	Fish-experimental	1x	2000	68846	48985	330
53	Fish-experimental	1x	2000	90811	65584	193
54	Fish-experimental	1x	2000	74787	49830	576
55	Fish-experimental	4x	1600	123766	81756	217
56	Fish-experimental	4x	1600	134588	91775	352
57	Fish-experimental	4x	1600	84101	53578	231
58	Fish-experimental	0.33x	1600	116890	82715	254
59	Fish-experimental	0.33x	1600	84523	60743	448
60	Fish-experimental	0.33x	1600	110206	79817	99
61	Fish-experimental	1x	1600	79386	57136	376
62	Fish-experimental	1x	1600	64949	40735	117

63	Fish-experimental	1x	1600	69483	52442	201
64	Fish-experimental	4x	1400	80657	53361	364

Table C.1 continued

Sample #	Sample type	Treatment	Time	Raw # of reads	Deblur reads	Unique reads
65	Fish-experimental	4x	1400	79794	59515	198
66	Fish-experimental	4x	1400	115980	82810	260
67	Fish-experimental	0.33x	1400	99910	64516	338
68	Fish-experimental	0.33x	1400	119239	83574	432
69	Fish-experimental	0.33x	1400	115788	87583	271
70	Fish-experimental	1x	1400	121103	85696	314
71	Fish-experimental	1x	1400	106963	74489	364
72	Fish-experimental	1x	1400	99058	62657	507
73	Fish-experimental	4x	1230	96336	66379	348
74	Fish-experimental	4x	1230	84904	54685	340
75	Fish-experimental	4x	1230	83820	56979	248
76	Fish-experimental	0.33x	1230	101830	68300	355
77	Fish-experimental	0.33x	1230	130553	87699	511
78	Fish-experimental	0.33x	1230	135543	90691	514
79	Fish-experimental	1x	1230	160706	105782	504
80	Fish-experimental	1x	1230	159196	107413	722
81	Fish-experimental	1x	1230	117745	76608	478
82	Fish-experimental	4x	1100	124295	95808	65
83	Fish-experimental	4x	1100	96310	58113	126
84	Fish-experimental	4x	1100	48041	37915	49
85	Fish-experimental	0.33x	1100	84565	60426	116
86	Fish-experimental	0.33x	1100	66028	44302	164
87	Fish-experimental	0.33x	1100	49018	24775	430
88	Fish-experimental	1x	1100	106353	63575	108
89	Fish-experimental	1x	1100	111881	50506	364
90	Fish-experimental	1x	1100	88993	66650	92
91	Water sample	N/A	N/A	113568	77514	421
92	Water sample	N/A	N/A	40640	28843	226
93	Water sample	N/A	N/A	109090	76378	377
94	Water sample	N/A	N/A	85256	57250	271
95	Water sample	N/A	N/A	171469	115199	390
96	Water sample	N/A	N/A	365098	239836	527
97	Fish at T0	N/A	N/A	261859	196804	47
98	Fish at T0	N/A	N/A	69591	53118	48
99	Fish at T0	N/A	N/A	76515	55345	263
100	Fish at T0	N/A	N/A	74552	51009	189
101	Fish at T0	N/A	N/A	135164	105414	44
102	Fish at T0	N/A	N/A	170804	103557	271
103	Food sample	N/A	N/A	15100	10321	216
104	Food sample	N/A	N/A	15837	11052	213
105	Food sample	N/A	N/A	20455	14374	238

Table C.2 A total of 57 ESVs showed significant ($p<0.05$) change in abundance over the two-day diel sampling period. Significant ESVs were identified using Analysis of Composition of Microbiomes (ANCOM). In ANCOM, higher W values are associated with greater change in abundance.

Taxon-ID	Reject null hypothesis	W
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Vibrionales;f__Vibrionaceae;g__Photobacterium	TRUE	579
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rickettsiales;f__mitochondria;__	TRUE	576
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Vibrionales;f__Vibrionaceae;__	TRUE	573
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Alteromonadales;f__Moritellaceae;g__Moritella	TRUE	570
k__Bacteria;p__Fusobacteria;c__Fusobacteriia;o__Fusobacteriales;f__Fusobacteriaceae;g__Psychrilyobacter	TRUE	575
k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Campylobacteraceae;g__Arcobacter	TRUE	575
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Paracoccus	TRUE	575
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Alteromonadales;f__Psychromonadaceae;g__Psychromonas	TRUE	573
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae;g__Staphylococcus	TRUE	551
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Leuconostocaceae;g__	TRUE	575
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus	TRUE	573
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Virgibacillus	TRUE	579
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;__;__	TRUE	576
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Psychrobacter	TRUE	578
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Planococcaceae;g__Sporosarcina	TRUE	576
k__Bacteria;p__Fusobacteria;c__Fusobacteriia;o__Fusobacteriales;f__Fusobacteriaceae;g__Cetobacterium	TRUE	570
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus	TRUE	525
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus	TRUE	575
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus	TRUE	569
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Geobacillus	TRUE	563
k__Bacteria;p__[Thermi];c__Deinococci;o__Thermales;f__Thermaceae;g__Thermus	TRUE	568
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__[Marinicellales];f__[Marinicellaceae];g__	TRUE	569
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Nocardiopsaceae;g__	TRUE	556
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__	TRUE	562
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;__;__	TRUE	558
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Hydrogenophaga	TRUE	570
k__Bacteria;p__Bacteroidetes;c__[Rhodothermi];o__[Rhodothermales];f__[Balneolaceae];g__	TRUE	579
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Alteromonadales;f__;g__	TRUE	558
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Planococcaceae;g__Ureibacillus	TRUE	567
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Citrobacter	TRUE	564
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;__	TRUE	557
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Rhodocyclales;f__Rhodocyclaceae;g__Azoarcus	TRUE	565
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Alteromonadales;f__Idiomarinaceae;g__Pseudidiomarina	TRUE	550
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__[Tissierellaceae];g__Tepidimicrobium	TRUE	552
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__;g__	TRUE	550
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;__;__	TRUE	566
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Rhodobaca	TRUE	567
k__Bacteria;p__Verrucomicrobia;c__Opitutae;o__;f__;g__	TRUE	548
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Planococcaceae;__	TRUE	539
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae;__	TRUE	522
k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Cryomorphaceae;g__	TRUE	518
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Peptostreptococcaceae;g__	TRUE	542
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella	TRUE	523
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;__	TRUE	527
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__	TRUE	530
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Proteus	TRUE	535
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Anoxybacillus	TRUE	545
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__[Tissierellaceae];g__Sporanaerobacter	TRUE	525
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Peptostreptococcaceae;g__Peptostreptococcus	TRUE	514
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Alteromonadales;f__[Chromatiaceae];g__	TRUE	512
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Enterococcaceae;g__Vagococcus	TRUE	528
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Leuconostocaceae;g__Leuconostoc	TRUE	506
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae;g__Jeotgalicoccus	TRUE	497
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Erwinia	TRUE	527
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__	TRUE	493
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Beijerinckiaceae;g__	TRUE	506
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Yersinia	TRUE	508

Table C.3. 111 KEGG pathways from PICRUST-generated metagenomes are predicted to show differential expression between fed (n=54) and unfed (n=34) states according to DESEQ2. Shown here is the DESEQ2 output.

Predicted Pathway	baseMean	log2 fold change unfed to fed	lfcSE	stat	pvalue	padj
Aldosteronearegulatedasodiumareabsorption	4.94352785	-3.050570459	0.44498258	-6.8554828	7.1072E-12	7.8473E-10
Pancreaticasecretion	25.5679128	-2.725841967	0.40408795	-6.745665	1.5233E-11	7.8473E-10
Glycosaminoglycanabiosynthesisachondroitinasulfate	41.0220647	-2.558859871	0.38813252	-6.592748	4.3176E-11	1.4104E-09
Cytokineareceptors	1.77354165	-2.526487549	0.54809381	-4.6095896	4.0346E-06	4.2364E-05
Cytokineacytokineareceptorainteraction	1.77354165	-2.526487549	0.54809381	-4.6095896	4.0346E-06	4.2364E-05
Shigellosis	1.01224545	-2.230519591	0.56963109	-3.9157266	9.0132E-05	0.00055206
BiosynthesisaofatypeaIIapolyketideabackbone	9.22243004	-2.209711301	0.61107327	-3.6161151	0.00029906	0.00165892
Colorectalacancer	590.771482	-2.018897948	0.29909317	-6.7500636	1.4778E-11	7.8473E-10
Toxoplasmosis	590.771482	-2.018897948	0.29909317	-6.7500636	1.4778E-11	7.8473E-10
Viralamyocarditis	591.197605	-2.017846212	0.29907233	-6.7470174	1.5092E-11	7.8473E-10
InfluenzaaA	591.625133	-2.014564969	0.29906239	-6.7362698	1.625E-11	7.8473E-10
Smallacellalungacancer	593.094971	-2.004302466	0.29843897	-6.7159542	1.8684E-11	7.8473E-10
p53asignalingapathway	606.064021	-1.949013948	0.29442796	-6.6196633	3.6002E-11	1.3231E-09
Apoptosis	621.079628	-1.924548217	0.29716105	-6.4764485	9.3907E-11	2.7609E-09
Germination	205.296701	-1.706034068	0.42894686	-3.977262	6.9713E-05	0.00045546
Bacterialainvasionaofaepithelialacells	7.92772063	-1.355097248	0.36540331	-3.7084975	0.00020849	0.00122594
Parkinson'sadisease	3974.25956	-1.303156035	0.2620549	-4.9728359	6.5981E-07	1.021E-05
Glycosphingolipidabiosynthesisaganglioaseries	1522.2218	-1.270601046	0.28401218	-4.4737555	7.6858E-06	7.0613E-05
Cardiacamuscleacontraction	3382.78719	-1.198668806	0.26686323	-4.4916972	7.0658E-06	6.9154E-05
Staphylococcusaaureusainfection	479.714996	-1.14280605	0.34655688	-3.2976002	0.00097515	0.00455069
Vibrioacholeraeainfection	678.221433	-0.886421797	0.31040934	-2.8556544	0.00429482	0.0146823
Amoebiasis	157.248515	-0.86008259	0.29148276	-2.9507152	0.00317039	0.01226441
DaArginineaandaDaornithineametabolism	142.623307	-0.848280542	0.2921909	-2.9031724	0.00369403	0.01374218
Pertussis	1120.82021	-0.827486388	0.26555471	-3.1160674	0.0018328	0.00750622
Isoflavonoidabiosynthesis	7.94603985	-0.808499918	0.34028266	-2.3759657	0.01750309	0.0476473
Transcriptionarelatedaproteins	802.895323	-0.768639748	0.29490494	-2.6063983	0.00915	0.02802187
Vibrioacholeraeapathogenicacycle	24140.0427	-0.675743114	0.1525456	-4.4297778	9.433E-06	8.404E-05
Lipopolysaccharideabiosynthesis	17479.3157	-0.653936589	0.22584945	-2.8954536	0.00378611	0.01374218
Renalacellacarcinoma	1466.64623	-0.596367553	0.22344429	-2.6689765	0.00760828	0.02428972
Huntington'sadisease	8658.36819	-0.585304001	0.09344894	-6.2633559	3.7678E-10	1.007E-08
AmyotrophicalateralasclerosisaALSa	2212.49031	-0.571675981	0.24192689	-2.3630114	0.01812711	0.04889332
Ubiquinoneaandaotheraterpenoidaquinoneabiosynthesis	23859.0566	-0.550024518	0.22144382	-2.4838106	0.01299849	0.03710249
Lipopolysaccharideabiosynthesisaproteins	24719.1205	-0.546986375	0.15635965	-3.4982579	0.00046831	0.00245862
Flagellaraassembly	61091.1484	-0.533522322	0.15488827	-3.4445624	0.00057198	0.00289937
Alzheimer'sadisease	9814.41755	-0.495213976	0.09401755	-5.2672503	1.3848E-07	2.7143E-06
Cellamotilityaandasecretion	19237.5332	-0.40684918	0.09985568	-4.0743719	4.6139E-05	0.00033085
Membraneaandaintracellularastructuralamolecules	53061.0008	-0.398512399	0.09903194	-4.0240796	5.7199E-05	0.00039585
Glycanabiosynthesisaandametabolism	6852.70686	-0.388441658	0.13473043	-2.8831026	0.00393779	0.01398262
Bacterialamotilityaproteins	153509.677	-0.368652397	0.12840106	-2.8711008	0.00409045	0.01431658
Secretionasystem	176267.365	-0.36153145	0.08990584	-4.0212231	5.7897E-05	0.00039585
Bacterialachemotaxis	67072.4692	-0.351962284	0.12596625	-2.7940999	0.00520444	0.01758742
Bacterialasecretionasystem	70667.8596	-0.297668319	0.06442942	-4.6200681	3.8361E-06	4.2364E-05
Lipoicaacidametabolism	6581.59097	-0.256572527	0.04287106	-5.9847491	2.1672E-09	5.3097E-08
Pathwaysainacancer	6029.06717	-0.206358784	0.04217834	-4.8925295	9.9548E-07	1.3937E-05
Functionaunknown	170347.374	-0.199772629	0.0602173	-3.317529	0.00090817	0.00430651
Electronatransferacarriers	6072.23717	-0.191408708	0.07908066	-2.4204239	0.01550243	0.04299729
Twoacomponentasystem	208776.545	-0.185656896	0.05959727	-3.1151912	0.00183826	0.00750622
Nitrogenametabolism	82751.8577	-0.144175955	0.04545235	-3.1720246	0.0015138	0.00654497
Oxidativeaphosphorylation	118169.891	-0.142828818	0.02869199	-4.9780028	6.4244E-07	1.021E-05
Chaperonesaandafoldingacatalysts	96488.024	-0.13504805	0.05386972	-2.5069381	0.0121782	0.03580391
Otheraionacoupledatransporters	134288.61	-0.129841634	0.04465937	-2.9073771	0.00364474	0.01374218
CitrateacycleaTCAacyclea	65314.3826	-0.107497795	0.04474481	-2.4024638	0.01628505	0.04474583
Phenylalanineatyrosineaandatryptophanabiosynthesis	63180.1407	0.054360181	0.02118305	2.56621104	0.01028163	0.03084488
Folateabiosynthesis	41940.8903	0.071651125	0.02687921	2.66567108	0.00768348	0.02428972
Prenyltransferases	30833.7563	0.08300332	0.0253629	3.27262788	0.00106553	0.00486927
Carbonafixationapathwaysainaprokaryotes	113963.602	0.086625585	0.03005403	2.88232889	0.00394747	0.01398262
Isoquinolineaalkaloidabiosynthesis	5998.08503	0.098104986	0.04160136	2.35821579	0.01836302	0.04907933
Alanineaaspartateaandaglutamateametabolism	107076.707	0.100055437	0.02923172	3.42283794	0.00061971	0.00307605
TypeaIIadiabetesamellitus	4771.47451	0.10264161	0.03396125	3.02231577	0.00250849	0.00996615
Valinealeucineaandaisoleucineabiosynthesis	66208.4859	0.105572972	0.02117068	4.98675298	6.1402E-07	1.021E-05
Arginineaandaprolineametabolism	124815.938	0.106663899	0.01993223	5.35132798	8.7311E-08	1.9746E-06
MAPKasignalingapathwayayeast	4816.89648	0.108769539	0.03336574	3.25991658	0.00111445	0.00496437
Replicationarecombinationaandarepairaproteins	78319.9777	0.112908862	0.02713281	4.16134028	3.1639E-05	0.00023851
Generalafunctionapredictionaonly	393699.315	0.119407395	0.02662379	4.48498869	7.2918E-06	6.9154E-05
Energyametabolism	106874.86	0.120663884	0.03057176	3.94690643	7.9167E-05	0.00050598
Arachidonicaacidametabolism	5400.63519	0.132622985	0.03662511	3.62109415	0.00029336	0.00165861
Pyruvateametabolism	123629.73	0.133456184	0.04606534	2.89710627	0.00376622	0.01374218
Transcriptionamachinery	84178.9688	0.134643918	0.04827747	2.7889597	0.00528776	0.01766594
Carbohydrateametabolism	11190.4859	0.137953668	0.03454849	3.99304456	6.523E-05	0.00043586
Drugametabolismaotheraenzymes	33987.323	0.144005962	0.05786758	2.48854322	0.01282676	0.03697126
PantothenateaandaCoAabiosynthesis	58149.4005	0.147348396	0.04309282	3.41932576	0.00062777	0.00307605
Fattyaacidabiosynthesis	56572.1589	0.148020418	0.02884896	5.13087463	2.884E-07	5.2993E-06
Cysteineaandamethionineametabolism	97314.6598	0.150461557	0.05846911	2.57335114	0.0100719	0.03052719
Glycolysisa/aGluconeogenesis	123117.06	0.151410872	0.03551897	4.26281714	2.0187E-05	0.00016957
Insulinasignalingapathway	5225.08244	0.151460307	0.03231648	4.68678236	2.7753E-06	3.5184E-05
Porphyrinaandachlorophyllametabolism	127531.406	0.155345881	0.05660961	2.74416114	0.00606658	0.0195997
Thiamineametabolism	45144.8472	0.165825342	0.06572445	2.52303888	0.01163455	0.0345511
Nucleotideaexcisionarepair	38554.4421	0.166821164	0.05562022	2.99928996	0.0027061	0.0106079
Baseaexcisionarepair	47448.6264	0.17167225	0.03951496	4.34448759	1.396E-05	0.00012071
Methaneametabolism	127029.472	0.17763508	0.03675805	4.83254948	1.348E-06	1.8014E-05
Others	136906.732	0.184012593	0.06419653	2.8663949	0.00415176	0.0143602
Aminoaacidametabolism	26618.2296	0.189546545	0.04503433	4.20893459	2.5658E-05	0.00020388

Tyrosineametabolism	52023.6028	0.195950698	0.07408801	2.64483691	0.00817303	0.02556246
Ascorbateaandaaldarateametabolism	23332.9496	0.19699723	0.07884303	2.49860029	0.01246849	0.0362944

Table C.3 continued

Predicted Pathway	**baseMean**	**log2 fold change unfed to fed**	**lfcSE**	**stat**	**pvalue**	**padj**
Proteinaprocessingainaendoplasmicareticulum	5118.44052	0.201703509	0.04455278	4.52729339	5.9744E-06	6.0568E-05
Tropaneapiperidineaandapyridineaalkaloidabiosynthesis	15389.5264	0.216797277	0.05512223	3.93302802	8.3882E-05	0.00052471
Starchaandasucroseametabolism	73273.6459	0.219849911	0.08001332	2.74766654	0.0060021	0.0195997
Streptomycinabiosynthesis	18684.0421	0.222899325	0.06221947	3.58246907	0.00034036	0.00181939
Glutamatergicasynapse	10647.7407	0.227467056	0.05497923	4.13732702	3.5138E-05	0.00025826
Histidineametabolism	73386.4135	0.233809953	0.04760914	4.91103113	9.0599E-07	1.3318E-05
Cyanoaminoaacidametabolism	22608.8467	0.237268069	0.04495897	5.27743574	1.31E-07	2.7143E-06
Phenylpropanoidabiosynthesis	8743.34875	0.237477168	0.07476919	3.17613687	0.00149251	0.00654497
Proximalatubuleabicarbonateareclamation	3833.76986	0.250692031	0.08033455	3.12060031	0.00180483	0.00750622
Polyketideasugaraunitabiosynthesis	7655.57574	0.275312851	0.09996324	2.75414101	0.00588464	0.01943915
Linoleicaacidametabolism	6619.89702	0.277676753	0.09557381	2.90536462	0.00366826	0.01374218
Terpenoidabackboneabiosynthesis	61703.6983	0.277724842	0.05934598	4.67975867	2.8721E-06	3.5184E-05
Glycerolipidametabolism	51470.1418	0.27842748	0.06560612	4.24392533	2.1964E-05	0.00017938
Proteasome	5248.37489	0.326025678	0.08854206	3.68215584	0.00023127	0.0013332
Biosynthesisaofavancomycinagroupaantibiotics	4079.0159	0.332809031	0.07173036	4.6397235	3.4888E-06	4.1028E-05
Sphingolipidametabolism	10618.3867	0.388282639	0.10408417	3.73046776	0.00019112	0.00114675
Butirosinaandaneomycinabiosynthesis	3461.32498	0.416855742	0.11588366	3.59719169	0.00032167	0.00175132
Xyleneadegradation	10752.8133	0.419641249	0.13420277	3.12691936	0.00176648	0.00750622
Dioxinadegradation	11614.0793	0.439385406	0.13163381	3.33793731	0.00084403	0.00406794
Synthesisaandadegradationaofaketoneabodies	14883.7389	0.509164797	0.20888093	2.43758389	0.01478578	0.04140019
Primaryabileaacidabiosynthesis	5268.6358	0.516362407	0.14981701	3.44662063	0.00056765	0.00289937
Benzoateadegradation	71245.7212	0.547419748	0.20764835	2.63628273	0.00838199	0.02594004
Polycyclicaaromaticahydrocarbonadegradation	27930.388	0.696133626	0.22847258	3.04690232	0.00231213	0.00931186
Bisphenoladegradation	23627.0421	0.720787423	0.29498043	2.44350931	0.01454519	0.04111815
Betalainabiosynthesis	345.238343	1.04621745	0.44555741	2.34810918	0.01886899	0.04997732
Vasopressinaregulatedawaterareabsorption	2.94276364	1.641162389	0.50192745	3.26972033	0.00107654	0.00486927
Melanogenesis	14.5317296	2.655293221	0.63254706	4.19777971	2.6954E-05	0.00020854

Table C.4 Sequencing statistics and sample information for transcriptomes following BLAST against the SEED database.

Sample #	Time	Feeding status	Total reads	Bacterial read	Bacterial with SEED hits
RNA1	1100	unfed	86,889	40,650	12,959
RNA2	1100	unfed	155,396	39,992	11,481
RNA3	1100	unfed	58,611	29,113	6,251
RNA4	1100	unfed	60,567	28,692	7,430
RNA5	1100	unfed	180,661	32,022	7,354
RNA6	1230	fed	681,223	228,476	64,808
RNA7	1230	fed	243,191	69,363	18,121
RNA8	1230	fed	78,447	23,886	5,531
RNA9	1230	fed	325,803	83,532	21,480
RNA10	1230	fed	318,117	46,764	10,892
		Totals	**2,188,905**	**622,490**	**166,307**

Table C.5. Functional gene expression differs in the fed and unfed states. Shown here are all SEED genes showing differential expression between the fed (1230, n=5) and unfed (1100, n=5) state identified using DESEQ2 in R ($p<0.05$). Shown here is the DESEQ2 output.

Gene	baseMean	log2 change unfed to fed	lfcSE	stat	pvalue	padj
CRISPR-associated helicase Cas3	3.28	-4.89	1.43	-3.42	0.00	0.04
Late competence protein ComGA, access of DNA to ComEA	2.34	-4.54	1.36	-3.33	0.00	0.05
CRISPR-associated protein Cas2	2.35	-4.51	1.42	-3.18	0.00	0.07
Two-component response regulator yesN, associated with MetSO reductase	1.98	-4.27	1.45	-2.95	0.00	0.10
Anaerobic sulfite reductase subunit C (EC 1.8.1.-)	1.99	-4.14	1.51	-2.74	0.01	0.14
Ribosome-associated heat shock protein	1.80	-4.09	1.52	-2.69	0.01	0.15
Periplasmic thiol:disulfide interchange protein DsbA	3.19	-4.08	1.28	-3.18	0.00	0.07
ABC transporter membrane-spanning permease, Pep export, Vex3	3.09	-4.07	1.26	-3.24	0.00	0.06
FIG138315: Putative alpha helix protein	1.76	-4.02	1.70	-2.37	0.02	0.21
Substrate-specific component BL0695 of predicted ECF transporter	2.90	-3.99	1.32	-3.01	0.00	0.09
Transcriptional repressor for NAD biosynthesis in gram-positives	1.49	-3.89	1.64	-2.37	0.02	NA
YoeB toxin protein	1.56	-3.86	1.56	-2.48	0.01	NA
FIG009886: phosphoesterase	1.57	-3.81	1.60	-2.38	0.02	NA
Alpha-D-GlcNAc alpha-1,2-L-rhamnosyltransferase (EC 2.4.1.-)	1.45	-3.80	1.71	-2.23	0.03	NA
Hcp transcriptional regulator HcpR (Crp/Fnr family)	1.44	-3.71	1.74	-2.13	0.03	NA
archaeal ATPase, fused to C-terminal DUF234 domain	1.47	-3.70	1.61	-2.31	0.02	NA
Cobalt-precorrin-8x methylmutase (EC 5.4.1.2)	2.12	-3.61	1.41	-2.56	0.01	0.18
Alpha-glycerophosphate oxidase (EC 1.1.3.21)	1.26	-3.61	1.74	-2.07	0.04	NA
Methylthioadenosine deaminase	2.07	-3.56	1.37	-2.60	0.01	0.17
Membrane protein CarB	1.25	-3.55	1.69	-2.10	0.04	NA
Spermidine synthase (EC 2.5.1.16)	2.19	-3.54	1.39	-2.54	0.01	0.18
Two-component system histidine kinase	1.25	-3.51	1.74	-2.02	0.04	NA
Duplicated ATPase component BL0693	2.04	-3.51	1.46	-2.41	0.02	0.20
Putative mobilization protein BF0133	2.14	-3.50	1.39	-2.51	0.01	0.18
ATP synthase F0 sector subunit b	3.12	-3.44	1.33	-2.58	0.01	0.17
Negative regulator of genetic competence MecA	2.12	-3.38	1.40	-2.41	0.02	0.20
Transcriptional regulator KdgR, KDG operon repressor	1.81	-3.35	1.52	-2.21	0.03	0.25
Branched-chain amino acid transport ATP-binding protein LivF (TC 3.A.1.4.1)	2.87	-3.35	1.24	-2.70	0.01	0.15
Cob(I)alamin adenosyltransferase PduO (EC 2.5.1.17)	1.74	-3.24	1.44	-2.25	0.02	0.23
Multiple sugar ABC transporter, membrane-spanning permease protein MsmG	1.63	-3.23	1.60	-2.02	0.04	NA
TPR domain protein, putative component of TonB system	1.90	-3.17	1.47	-2.15	0.03	0.26
Substrate-specific component RibU of riboflavin ECF transporter	1.76	-3.15	1.55	-2.04	0.04	0.30
Histidine kinase of the competence regulon ComD	2.59	-3.12	1.23	-2.53	0.01	0.18
Putative DNA-binding protein in cluster with Type I restriction-modification system	2.67	-3.11	1.32	-2.36	0.02	0.21
Glycogen branching enzyme, GH-57-type, archaeal (EC 2.4.1.18)	1.73	-3.09	1.52	-2.04	0.04	0.30
Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.7)	2.42	-3.01	1.31	-2.29	0.02	0.23
Hypothetical similar to thiamin biosynthesis lipoprotein ApbE	1.53	-2.98	1.50	-1.99	0.05	NA
Choline kinase (EC 2.7.1.32)	2.19	-2.97	1.40	-2.13	0.03	0.27
Pullulanase (EC 3.2.1.41)	5.82	-2.95	0.84	-3.50	0.00	0.04
CRISPR-associated protein, Csn1 family	5.09	-2.93	0.93	-3.16	0.00	0.07
Serine endopeptidase ScpC (EC 3.4.21.-)	2.20	-2.91	1.29	-2.26	0.02	0.23
Serine protease precursor MucD/AlgY associated with sigma factor RpoE	2.35	-2.86	1.43	-1.99	0.05	0.32
L-threonine 3-O-phosphate decarboxylase (EC 4.1.1.81)	3.17	-2.86	1.26	-2.27	0.02	0.23
NAD-specific glutamate dehydrogenase (EC 1.4.1.2)	3.75	-2.79	1.08	-2.57	0.01	0.18
RNA methyltransferase, TrmA family	2.15	-2.75	1.32	-2.08	0.04	0.29
L-serine dehydratase, beta subunit (EC 4.3.1.17)	3.07	-2.75	1.31	-2.10	0.04	0.28
Aldose 1-epimerase (EC 5.1.3.3)	3.03	-2.67	1.19	-2.25	0.02	0.23
CRISPR-associated protein, Cse4 family	2.62	-2.62	1.20	-2.19	0.03	0.26
HPr kinase/phosphorylase (EC 2.7.1.-) (EC 2.7.4.-)	2.65	-2.45	1.19	-2.05	0.04	0.30
Uracil permease	6.30	-2.43	0.77	-3.16	0.00	0.07
RNA-binding protein Jag	3.50	-2.42	1.07	-2.26	0.02	0.23
Flavodoxin	3.42	-2.40	1.04	-2.30	0.02	0.23
Fumarate reductase flavoprotein subunit (EC 1.3.99.1)	5.17	-2.38	0.88	-2.69	0.01	0.15
Sucrose operon repressor ScrR, LacI family	3.72	-2.34	1.00	-2.34	0.02	0.22
Chromosome partition protein smc	12.35	-2.33	0.54	-4.32	0.00	0.00
Lipopolysaccharide heptosyltransferase III (EC 2.4.1.-)	5.33	-2.29	0.83	-2.77	0.01	0.14
tRNA-specific adenosine-34 deaminase (EC 3.5.4.-)	3.52	-2.23	1.08	-2.06	0.04	0.29
Multiple sugar ABC transporter, ATP-binding protein	3.88	-2.06	0.96	-2.16	0.03	0.26
Flagellar biosynthesis protein FliC	3.92	-1.97	0.88	-2.25	0.02	0.23
Phosphopantothenoylcysteine decarboxylase (EC 4.1.1.36)	3.64	-1.95	0.94	-2.07	0.04	0.29
Galactokinase (EC 2.7.1.6)	6.41	-1.94	0.77	-2.52	0.01	0.18
Heat-inducible transcription repressor HrcA	5.10	-1.88	0.83	-2.26	0.02	0.23
Replicative DNA helicase (EC 3.6.1.-) [SA14-24]	5.19	-1.87	0.83	-2.25	0.02	0.23
DNA primase (EC 2.7.7.-)	12.60	-1.86	0.47	-3.93	0.00	0.01
Dihydroorotate dehydrogenase, catalytic subunit (EC 1.3.3.1)	4.73	-1.83	0.83	-2.20	0.03	0.25
Sucrose-6-phosphate hydrolase (EC 3.2.1.B3)	6.76	-1.80	0.66	-2.72	0.01	0.14
Beta-fimbriae probable major subunit	4.76	-1.79	0.86	-2.07	0.04	0.29
Potassium uptake protein TrkH	7.80	-1.76	0.65	-2.72	0.01	0.14
Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7)	4.62	-1.75	0.83	-2.11	0.03	0.28
Galactose/methyl galactoside ABC transport system	4.67	-1.73	0.86	-2.00	0.05	0.31
PF00070 family, FAD-dependent NAD(P)-disulphide oxidoreductase	4.65	-1.70	0.83	-2.05	0.04	0.30
Choline binding protein A	4.39	-1.70	0.84	-2.02	0.04	0.30
Transcriptional activator of cad operon	11.07	-1.68	0.67	-2.49	0.01	0.18
DNA repair protein RecN	7.50	-1.67	0.67	-2.50	0.01	0.18
Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94)	8.46	-1.64	0.58	-2.81	0.00	0.13

Table C.5 continued

Gene	baseMean	log2 change unfed to fed	lfcSE	stat	pvalue	padj
Enoyl-[acyl-carrier-protein] reductase [FMN] (EC 1.3.1.9)	6.14	-1.63	0.70	-2.31	0.02	0.23
Recombination inhibitory protein MutS2	7.00	-1.59	0.61	-2.63	0.01	0.16
DNA recombination protein RmuC	7.03	-1.58	0.70	-2.25	0.02	0.23
1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)	14.85	-1.58	0.44	-3.58	0.00	0.03
GTP-binding protein EngB	5.87	-1.57	0.70	-2.25	0.02	0.23
tRNA(Ile)-lysidine synthetase (EC 6.3.4.19)	7.39	-1.56	0.65	-2.40	0.02	0.20
Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)	7.05	-1.55	0.63	-2.47	0.01	0.19
Pyruvate formate-lyase (EC 2.3.1.54)	18.29	-1.35	0.37	-3.64	0.00	0.03
Pyruvate,phosphate dikinase (EC 2.7.9.1)	8.47	-1.35	0.55	-2.48	0.01	0.19
DNA polymerase III subunits gamma and tau (EC 2.7.7.7)	10.50	-1.29	0.51	-2.52	0.01	0.18
SSU ribosomal protein S3p (S3e)	7.17	-1.27	0.58	-2.20	0.03	0.25
Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8)	10.92	-1.23	0.50	-2.44	0.01	0.20
Acetate kinase (EC 2.7.2.1)	11.11	-1.13	0.47	-2.41	0.02	0.20
CRISPR-associated helicase Cas3, protein	8.64	-1.11	0.54	-2.05	0.04	0.30
Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	11.74	-1.07	0.49	-2.20	0.03	0.25
Multi antimicrobial extrusion protein	15.00	-1.04	0.40	-2.59	0.01	0.17
Phosphoglycerate kinase (EC 2.7.2.3)	10.50	-1.03	0.48	-2.16	0.03	0.26
Adenylosuccinate synthetase (EC 6.3.4.4)	12.95	-1.02	0.44	-2.32	0.02	0.22
Beta-galactosidase (EC 3.2.1.23)	15.79	-1.02	0.40	-2.53	0.01	0.18
DNA recombination and repair protein RecF	14.36	-1.01	0.39	-2.56	0.01	0.18
Phosphoglucosamine mutase (EC 5.4.2.10)	11.49	-1.01	0.50	-2.03	0.04	0.30
UDP-glucose 4-epimerase (EC 5.1.3.2)	13.08	-1.00	0.44	-2.28	0.02	0.23
DNA mismatch repair protein MutS	16.86	-0.98	0.42	-2.32	0.02	0.22
Trk system potassium uptake protein TrkA	11.95	-0.97	0.45	-2.15	0.03	0.26
Pyruvate-flavodoxin oxidoreductase (EC 1.2.7.-)	12.46	-0.96	0.43	-2.27	0.02	0.23
Chromosomal replication initiator protein DnaA	12.71	-0.93	0.44	-2.10	0.04	0.29
Ribonucleotide reductase of class III (anaerobic), large subunit (EC 1.17.4.2)	14.96	-0.90	0.44	-2.06	0.04	0.29
Phosphoribosylformylglycinamidine synthase	16.48	-0.88	0.39	-2.27	0.02	0.23
DNA mismatch repair protein MutL	12.78	-0.88	0.42	-2.09	0.04	0.29
DNA polymerase III alpha subunit (EC 2.7.7.7)	34.78	-0.83	0.26	-3.20	0.00	0.07
Alanyl-tRNA synthetase (EC 6.1.1.7)	19.92	-0.82	0.35	-2.35	0.02	0.21
ATP-dependent DNA helicase UvrD/PcrA	19.93	-0.77	0.34	-2.25	0.02	0.23
Transcriptional regulation	39.47	-0.66	0.27	-2.43	0.02	0.20
Chaperone protein DnaK	23.11	-0.64	0.32	-1.99	0.05	0.32
DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)	36.19	-0.56	0.28	-1.99	0.05	0.32
Cell Division and Cell Cycle	287.30	-0.36	0.11	-3.34	0.00	0.05
DNA Metabolism	908.63	-0.34	0.05	-6.38	0.00	0.00
Protein Metabolism	1188.64	-0.20	0.05	-3.93	0.00	0.01
RNA Metabolism	691.76	-0.18	0.07	-2.72	0.01	0.14
Cofactors, Vitamins, Prosthetic Groups, Pigments	1514.07	0.16	0.05	3.16	0.00	0.07
Miscellaneous	816.50	0.16	0.07	2.35	0.02	0.21
Stress Response	622.89	0.18	0.09	2.07	0.04	0.29
Fatty Acids, Lipids, and Isoprenoids	439.90	0.19	0.09	2.16	0.03	0.26
Carbohydrates	1978.71	0.21	0.05	4.45	0.00	0.00
Metabolite damage and its repair or mitigation	329.91	0.22	0.09	2.36	0.02	0.21
Sulfur Metabolism	218.32	0.26	0.12	2.15	0.03	0.26
Amino Acids and Derivatives	1459.23	0.30	0.04	6.90	0.00	0.00
Phages, Prophages, Transposable elements	76.35	0.41	0.20	2.03	0.04	0.30
Respiration	498.87	0.45	0.08	5.43	0.00	0.00
Phosphorus Metabolism	153.80	0.55	0.15	3.72	0.00	0.02
Nitrogen Metabolism	204.03	0.65	0.12	5.32	0.00	0.00
Metabolism of Aromatic Compounds	193.38	0.84	0.15	5.75	0.00	0.00
Polyphosphate kinase (EC 2.7.4.1)	16.16	0.85	0.43	1.98	0.05	0.32
RND efflux system, inner membrane transporter CmeB	20.21	0.87	0.40	2.20	0.03	0.25
Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	53.43	0.90	0.22	4.14	0.00	0.01
Respiratory nitrate reductase alpha chain (EC 1.7.99.4)	12.81	0.93	0.43	2.15	0.03	0.26
Cobalt-zinc-cadmium resistance protein CzcA	21.29	0.93	0.37	2.53	0.01	0.18
Catalase (EC 1.11.1.6)	9.73	1.02	0.50	2.06	0.04	0.29
Plant cell walls and outer surfaces	10.11	1.05	0.50	2.08	0.04	0.29
Delta-1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12)	10.49	1.09	0.52	2.08	0.04	0.29
Membrane fusion component of tripartite multidrug resistance system	9.20	1.11	0.50	2.21	0.03	0.25
Lipoate synthase	6.74	1.23	0.62	1.99	0.05	0.32
Acetyl-coenzyme A synthetase (EC 6.2.1.1)	12.17	1.23	0.46	2.66	0.01	0.15
Exodeoxyribonuclease V gamma chain (EC 3.1.11.5)	6.65	1.27	0.64	1.99	0.05	0.32
Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)	8.61	1.31	0.54	2.41	0.02	0.20
Exodeoxyribonuclease V beta chain (EC 3.1.11.5)	7.74	1.36	0.54	2.50	0.01	0.18
Fumarate hydratase class II (EC 4.2.1.2)	6.76	1.37	0.67	2.04	0.04	0.30
Tyrosine-protein kinase Wzc (EC 2.7.10.2)	6.62	1.38	0.64	2.16	0.03	0.26
rRNA small subunit methyltransferase H	7.35	1.39	0.67	2.09	0.04	0.29
Dipeptide transport system permease protein DppC (TC 3.A.1.5.2)	7.46	1.40	0.60	2.34	0.02	0.22
NAD(P) transhydrogenase alpha subunit (EC 1.6.1.2)	5.44	1.43	0.69	2.08	0.04	0.29
D-serine/D-alanine/glycine transporter	13.46	1.46	0.42	3.49	0.00	0.04
Respiratory nitrate reductase beta chain (EC 1.7.99.4)	6.53	1.47	0.64	2.30	0.02	0.23
D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)	13.70	1.47	0.41	3.62	0.00	0.03
Pyruvate dehydrogenase E1 component (EC 1.2.4.1)	10.36	1.49	0.49	3.07	0.00	0.08
Phosphate transport system permease protein PstA (TC 3.A.1.7.1)	6.81	1.49	0.64	2.34	0.02	0.22
Acetylornithine deacetylase (EC 3.5.1.16)	4.77	1.49	0.75	2.00	0.05	0.31
Cobalt-zinc-cadmium resistance protein CzcD	9.62	1.50	0.56	2.69	0.01	0.15
NADH-ubiquinone oxidoreductase chain M (EC 1.6.5.3)	7.08	1.51	0.68	2.22	0.03	0.25
Glycerol-3-phosphate transporter	6.14	1.54	0.75	2.05	0.04	0.30
Gamma-glutamyltranspeptidase (EC 2.3.2.2)	5.56	1.55	0.70	2.20	0.03	0.25
Paraquat-inducible protein A	4.56	1.56	0.79	1.98	0.05	0.32

Table C.5 continued

Gene	baseMean	log2 change unfed to fed	lfcSE	stat	pvalue	padj
Alpha-glucosidase (EC 3.2.1.20)	7.32	1.65	0.64	2.60	0.01	0.17
Alanine dehydrogenase (EC 1.4.1.1)	13.48	1.67	0.48	3.47	0.00	0.04
Acyl-CoA dehydrogenase, short-chain specific (EC 1.3.99.2)	7.25	1.71	0.71	2.41	0.02	0.20
L-Proline/Glycine betaine transporter ProP	6.15	1.73	0.68	2.52	0.01	0.18
Methylmalonyl-CoA mutase (EC 5.4.99.2)	8.20	1.75	0.54	3.23	0.00	0.06
Glutamate synthase [NADPH] large chain (EC 1.4.1.13)	19.29	1.78	0.36	4.92	0.00	0.00
2-oxoglutarate dehydrogenase E1 component (EC 1.2.4.2)	10.73	1.82	0.48	3.75	0.00	0.02
Acetolactate synthase small subunit (EC 2.2.1.6)	5.48	1.82	0.73	2.49	0.01	0.18
Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)	4.70	1.82	0.78	2.34	0.02	0.22
4-keto-6-deoxy-N-Acetyl-D-hexosaminyl-(Lipid carrier) aminotransferase	4.21	1.85	0.94	1.98	0.05	0.32
Maltose phosphorylase (EC 2.4.1.8)	4.59	1.87	0.95	1.96	0.05	0.33
Lysine/cadaverine antiporter membrane protein CadB	4.17	1.96	0.91	2.15	0.03	0.26
Multicopper oxidase	4.72	1.98	0.80	2.47	0.01	0.19
Cytochrome O ubiquinol oxidase subunit I (EC 1.10.3.-)	3.86	2.01	0.94	2.15	0.03	0.26
Glutamate-ammonia-ligase adenylyltransferase (EC 2.7.7.42)	4.32	2.01	0.87	2.33	0.02	0.22
Acetylglutamate kinase (EC 2.7.2.8)	5.28	2.04	0.71	2.88	0.00	0.12
Myo-inositol 2-dehydrogenase (EC 1.1.1.18)	4.64	2.05	0.88	2.33	0.02	0.22
Succinyl-CoA ligase [ADP-forming] beta chain (EC 6.2.1.5)	5.16	2.07	0.78	2.66	0.01	0.15
Sulfite reductase [NADPH] hemoprotein beta-component (EC 1.8.1.2)	4.50	2.08	0.86	2.43	0.02	0.20
Aconitate hydratase 2 (EC 4.2.1.3)	7.20	2.09	0.65	3.20	0.00	0.07
Cytochrome c-type biogenesis protein DsbD	4.45	2.13	0.86	2.48	0.01	0.18
UDP-glucuronic acid oxidase	2.78	2.15	0.99	2.17	0.03	0.26
High-affinity choline uptake protein BetT	6.81	2.18	0.71	3.06	0.00	0.08
Thiazole biosynthesis protein ThiG	2.67	2.18	1.01	2.15	0.03	0.26
Phosphate transport system permease protein PstC (TC 3.A.1.7.1)	3.17	2.23	0.94	2.36	0.02	0.21
Nucleoside diphosphate kinase (EC 2.7.4.6)	2.61	2.23	1.04	2.13	0.03	0.27
Malate:quinone oxidoreductase (EC 1.1.5.4)	7.57	2.24	0.65	3.42	0.00	0.04
L-threonine transporter, anaerobically inducible	2.70	2.29	1.04	2.20	0.03	0.25
S-formylglutathione hydrolase (EC 3.1.2.12)	3.00	2.29	1.01	2.27	0.02	0.23
Dihydrolipoamide acetyltransferase	6.62	2.31	0.73	3.14	0.00	0.07
Menaquinone-specific isochorismate synthase (EC 5.4.4.2)	4.08	2.33	0.92	2.53	0.01	0.18
Choline dehydrogenase (EC 1.1.99.1)	4.59	2.41	0.95	2.53	0.01	0.18
Succinate-semialdehyde dehydrogenase [NADP+] (EC 1.2.1.79)	2.57	2.41	1.16	2.07	0.04	0.29
2-dehydro-3-deoxygluconate kinase (EC 2.7.1.45)	2.72	2.42	1.19	2.04	0.04	0.30
Glucose dehydrogenase, PQQ-dependent (EC 1.1.5.2)	2.49	2.42	1.23	1.97	0.05	0.32
2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol hydroxylase (EC 1.14.13.-)	2.44	2.43	1.23	1.98	0.05	0.32
D-amino acid dehydrogenase small subunit (EC 1.4.99.1)	4.63	2.46	0.88	2.81	0.01	0.13
Pantothenate:Na+ symporter (TC 2.A.21.1.1)	4.59	2.47	0.85	2.91	0.00	0.11
Transcriptional repressor of the lac operon	2.80	2.52	1.07	2.35	0.02	0.21
Succinate dehydrogenase hydrophobic membrane anchor protein	2.10	2.55	1.26	2.02	0.04	0.30
Flagellar hook-associated protein FlgL	2.83	2.57	1.17	2.20	0.03	0.25
Malate synthase G (EC 2.3.3.9)	3.18	2.59	1.21	2.15	0.03	0.26
1-phosphofructokinase (EC 2.7.1.56)	1.82	2.70	1.30	2.07	0.04	0.29
Co-activator of prophage gene expression IbrB	2.31	2.73	1.36	2.01	0.04	0.31
Uncharacterized protein YphG, TPR-domain containing	2.45	2.76	1.31	2.10	0.04	0.28
Poly(glycerol-phosphate) alpha-glucosyltransferase (EC 2.4.1.52)	3.29	2.78	1.17	2.37	0.02	0.21
LptA, protein essential for LPS transport across the periplasm	2.06	2.83	1.26	2.25	0.02	0.23
Hydroxymethylglutaryl-CoA reductase (EC 1.1.1.34)	2.64	2.83	1.13	2.51	0.01	0.18
Conserved protein YghR, with nucleoside triphosphate hydrolase domain	2.76	2.84	1.33	2.13	0.03	0.27
Methylcrotonyl-CoA carboxylase carboxyl transferase subunit (EC 6.4.1.4)	3.05	2.89	1.09	2.66	0.01	0.15
Nitrate/nitrite sensor protein (EC 2.7.3.-)	3.33	2.91	1.12	2.59	0.01	0.17
Uncharacterized siderophore biosynthesis protein near heme transporter HtsABC	3.18	2.92	1.23	2.37	0.02	0.21
Valine--pyruvate aminotransferase (EC 2.6.1.66)	4.25	2.95	0.91	3.23	0.00	0.06
LysR family transcriptional regulator YbhD	1.55	3.00	1.45	2.07	0.04	NA
Ribosomal-protein-L7p-serine acetyltransferase	2.59	3.04	1.21	2.51	0.01	0.18
Glycine cleavage system transcriptional activator GcvA	1.76	3.06	1.53	2.00	0.05	0.31
COG3178: Predicted phosphotransferase related to Ser/Thr protein kinases	1.94	3.07	1.47	2.09	0.04	0.29
Outer surface protein of unknown function, cellobiose operon	2.89	3.08	1.13	2.74	0.01	0.14
Co-activator of prophage gene expression IbrA	3.06	3.10	1.07	2.89	0.00	0.11
4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	2.20	3.14	1.32	2.37	0.02	0.21
2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60)	1.82	3.16	1.46	2.16	0.03	0.26
Molybdopterin biosynthesis protein MoeB	2.36	3.17	1.28	2.48	0.01	0.19
2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylic-acid synthase	1.75	3.20	1.41	2.27	0.02	0.23
Evolved beta-D-galactosidase, alpha subunit	1.98	3.22	1.31	2.46	0.01	0.19
Na(+) H(+) antiporter subunit D	2.50	3.22	1.21	2.65	0.01	0.15
L-fuculokinase (EC 2.7.1.51)	1.93	3.23	1.51	2.14	0.03	0.26
L-lactate dehydrogenase (EC 1.1.2.3)	3.87	3.25	1.10	2.97	0.00	0.10
Ferredoxin, 2Fe-2S	1.90	3.26	1.50	2.18	0.03	0.26
Pyridoxine 5'-phosphate synthase (EC 2.6.99.2)	1.94	3.27	1.55	2.11	0.04	0.28
Nitrite reductase [NAD(P)H] large subunit (EC 1.7.1.4)	5.16	3.31	0.96	3.44	0.00	0.04
Dolichol-phosphate mannosyltransferase (EC 2.4.1.83)	2.49	3.38	1.39	2.43	0.02	0.20
Membrane-bound lytic murein transglycosylase B precursor (EC 3.2.1.-)	1.34	3.40	1.72	1.98	0.05	NA
Catechol 1,2-dioxygenase (EC 1.13.11.1)	1.42	3.41	1.71	1.99	0.05	NA
Autotrophy	2.13	3.42	1.42	2.41	0.02	0.20
Nitrogen regulatory protein P-II	2.13	3.42	1.42	2.41	0.02	0.20
Hexuronate transporter	5.03	3.43	1.01	3.40	0.00	0.04
Hypoxanthine/guanine permease PbuG	1.40	3.45	1.72	2.01	0.04	NA
Salicylate hydroxylase (EC 1.14.13.1)	2.45	3.52	1.31	2.68	0.01	0.15
2-methylcitrate synthase (EC 2.3.3.5)	2.20	3.54	1.40	2.54	0.01	0.18
Homolog of fucose/glucose/galactose permeases	3.18	3.56	1.31	2.72	0.01	0.14
Glycerol-3-phosphate ABC transporter, permease protein UgpA (TC 3.A.1.1.3)	1.64	3.57	1.60	2.22	0.03	NA
Predicted cell-wall-anchored protein SasC (LPXTG motif)	5.69	3.59	0.98	3.67	0.00	0.03

Table C.5 continued

Gene	baseMean	log2 change unfed to fed	lfcSE	stat	pvalue	padj
Bacitracin export permease protein BceB	2.33	3.62	1.29	2.81	0.01	0.13
Trehalose-6-phosphate phosphatase (EC 3.1.3.12)	1.71	3.69	1.46	2.52	0.01	0.18
Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), permease component FhuB	1.68	3.76	1.54	2.44	0.01	NA
2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC 4.6.1.12)	2.64	3.76	1.22	3.09	0.00	0.08
Putrescine/proton symporter, putrescine/ornithine antiporter PotE	1.77	3.79	1.44	2.64	0.01	0.16
Molybdopterin-guanine dinucleotide biosynthesis protein MobA	3.04	3.80	1.22	3.12	0.00	0.07
RNA polymerase sigma factor RpoS	2.73	3.89	1.43	2.72	0.01	0.14
Copper resistance protein B	2.08	3.90	1.56	2.50	0.01	0.18
Thymidylate synthase thyX (EC 2.1.1.-)	2.08	4.00	1.49	2.68	0.01	0.15
DNA-binding heavy metal response regulator	1.97	4.01	1.47	2.73	0.01	0.14
Formate dehydrogenase N alpha subunit (EC 1.2.1.2)	2.22	4.13	1.46	2.83	0.00	0.13
Alkanesulfonates-binding protein	2.26	4.15	1.46	2.83	0.00	0.13
ATP-dependent RNA helicase RhlB	2.32	4.20	1.42	2.97	0.00	0.10
D-galactonate regulator, IclR family	2.31	4.21	1.50	2.80	0.01	0.13
Branched-chain alpha-keto acid dehydrogenase, E1 component (EC 1.2.4.4)	2.51	4.22	1.33	3.17	0.00	0.07
NADPH-dependent mycothiol reductase Mtr	2.45	4.24	1.60	2.66	0.01	0.15
putative Cytochrome bd2, subunit II	2.80	4.43	1.48	2.99	0.00	0.10
Molybdenum cofactor biosynthesis protein MoaE	1.75	4.50	1.58	2.85	0.00	0.12
Zinc transporter ZitB	1.62	4.50	1.71	2.64	0.01	NA
Pseudouridine 5'-phosphate glycosidase	1.74	4.51	1.65	2.73	0.01	0.14
Bifunctional autolysin Atl	5.10	4.54	1.18	3.83	0.00	0.02
PTS system, glucitol/sorbitol-specific IIB component	1.72	4.57	1.66	2.75	0.01	0.14
3-keto-L-gulonate-6-phosphate decarboxylase UlaD (EC 4.1.1.85)	1.86	4.60	1.60	2.87	0.00	0.12
Mgl repressor and galactose ultrainduction factor GalS	1.92	4.66	1.52	3.06	0.00	0.08
5-keto-D-gluconate 5-reductase (EC 1.1.1.69)	1.94	4.70	1.58	2.98	0.00	0.10
Enterobactin synthetase component F, serine activating enzyme (EC 2.7.7.-)	2.02	4.73	1.63	2.91	0.00	0.11
Zn(II) and Co(II) transmembrane diffusion facilitator	2.01	4.78	1.40	3.41	0.00	0.04

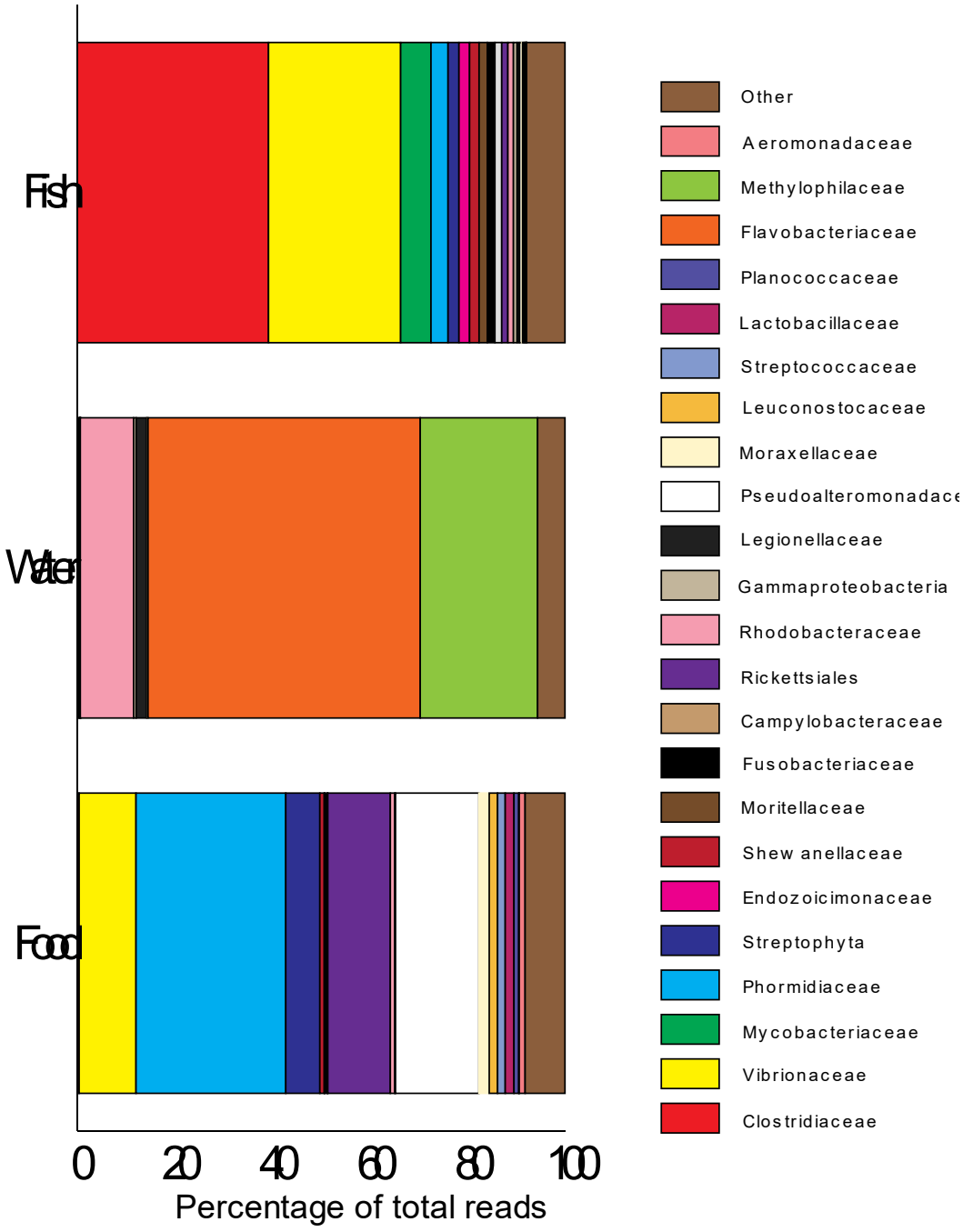

Figure C.1. Taxonomic composition varies according to sample type. Plotted are average percent relative abundances for dominant bacterial families in experimental fish (n=89), water (n=6), and food (n=3). Significant differences in community composition between sample types was confirmed using Bray-Curtis similarity and principal coordinate analysis ($p<0.05$).

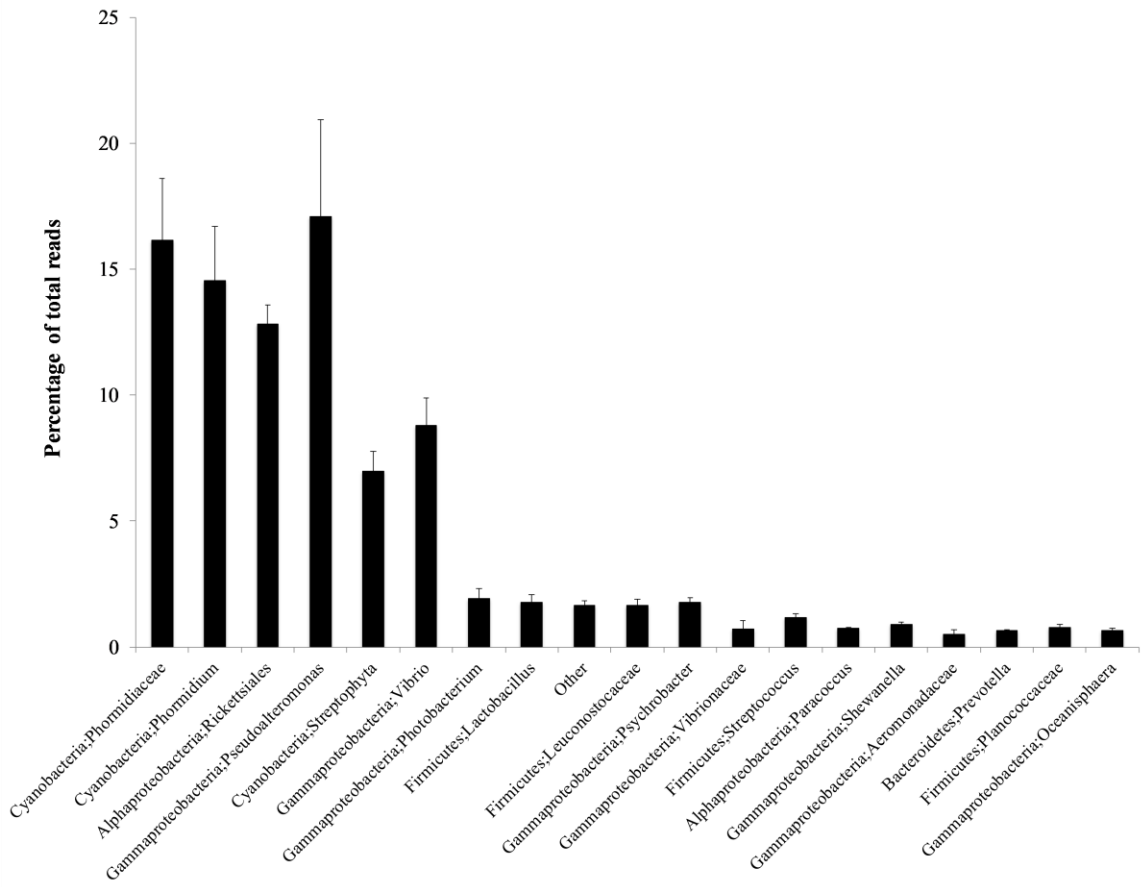

Figure C.2. Lowest taxonomic classification and average relative abundance of the top 18 ESVs detected in food samples (n=3)

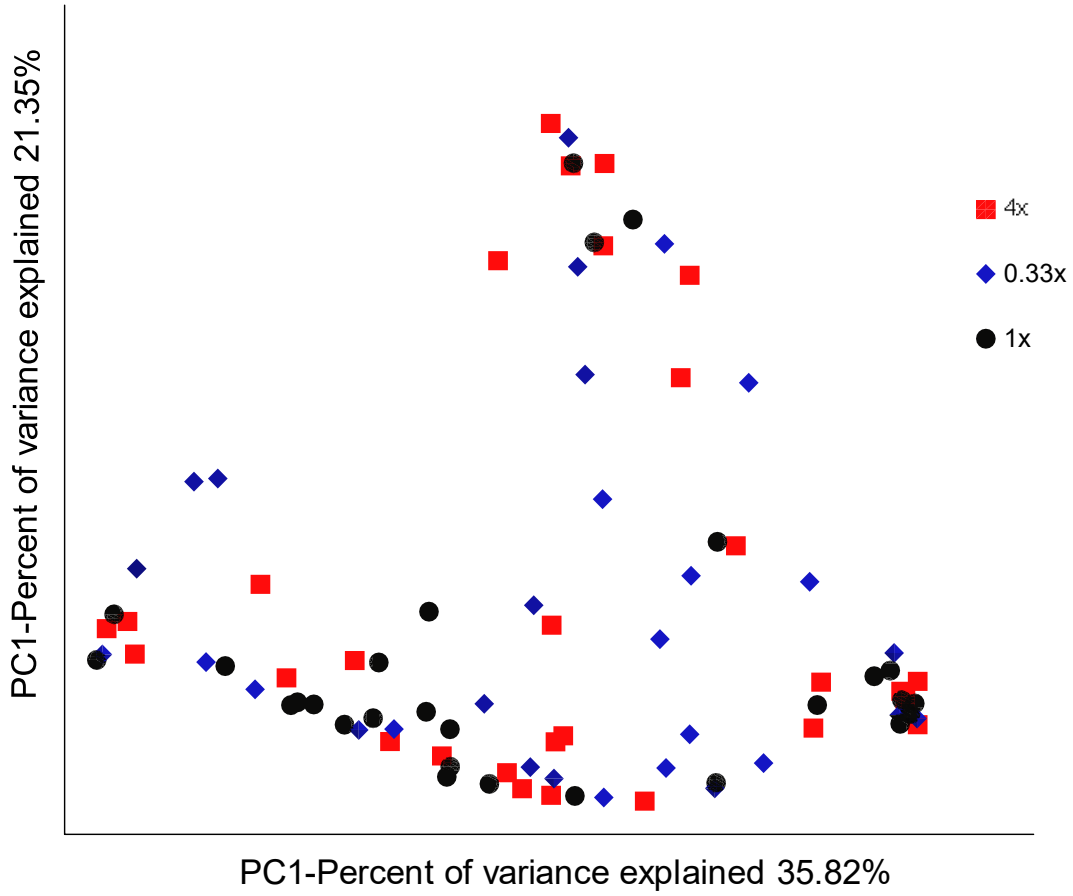

Figure C.3. There was no significant clustering of samples based on feeding frequency regime (p>0.05). Principal coordinate analysis based on Bray-Curtis similarity with all samples rarefied 20,562 sequences.

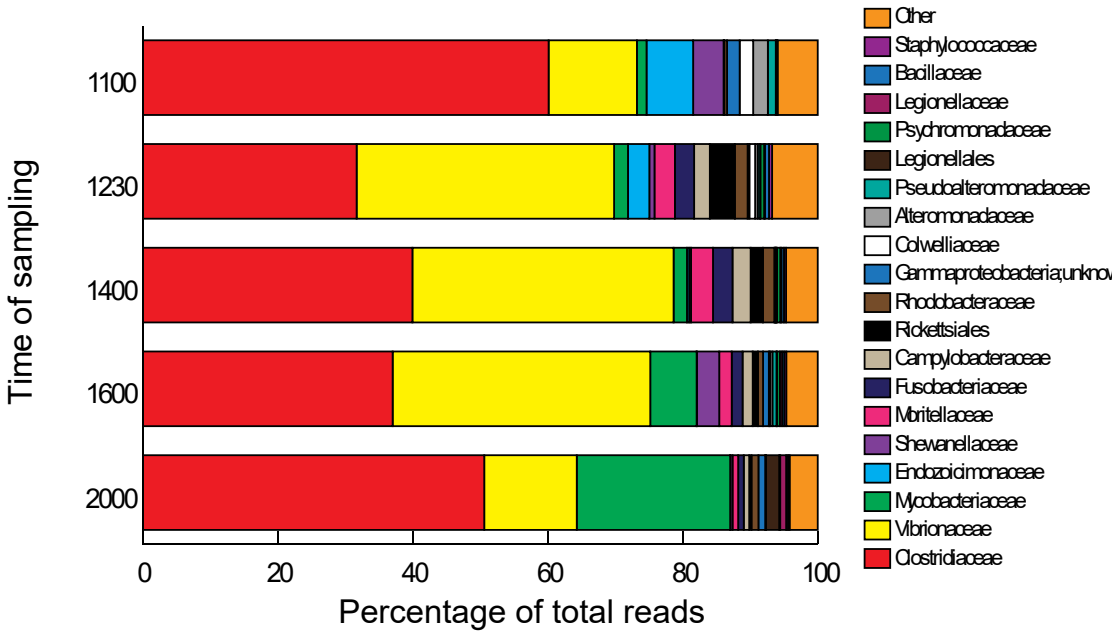

Figure C.4. Taxonomic composition varies according to sample type. Plotted are average percent relative abundances for dominant bacterial families in experimental fish (n=89), water (n=6), and food (n=3). Significant differences in community composition between sample types was confirmed using Bray-Curtis similarity and principal coordinate analysis ($p<0.05$).