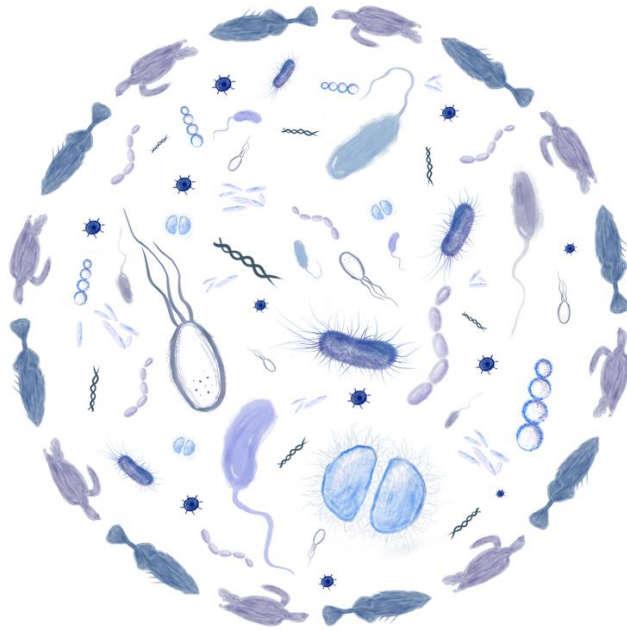# Host-microbe interactions in wild vertebrate populations

Adrienne E. Kerley

Submitted in partial fulfilment of the requirements of the Degree

of Doctor of Philosophy

School of Biological and Chemical Sciences

Queen Mary University of London

May 2021

I, Adrienne Kerley, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature:

Date: 25/05/2021

Details of Collaboration

Chapter 1: Adrienne Kerley (A.K) conducted writing. Christophe Eizaguirre (C.E) provided guidance and suggested edits.

Chapter 2: Naga Apoorva Kasala, Martin Kalbe (M.K), Manfred Milinski (M.M), and C.E designed and conducted the original sample collection. A.K and C.E designed specific study and hypotheses tested in this chapter. A.K conducted data analyses and writing. A.K and C.E interpreted the results. Comments and manuscript edits were provided by C.E, Lee Henry (L.H) and China Hanson (C.H).

Chapter 3: C.E and M.M conceived the original study design. A.K and C.E designed the specific study and hypotheses tested in this chapter. A.K, C.E, M.K, Christoph L Gahr (C.L.G), and Tina Henrich (T.H) conducted the fish collections, dissections, and parasite screening. C.L.G carried out sex-typing and genotyping. A.K carried out all stable isotope analysis. A.K conducted the data analyses and writing. A.K and C.E interpreted the results. Comments and manuscript edits by C.E, L.H and C.H.

Chapter 4: C.E and M.M conceived the original study design. A.K and C.E designed the specific study and hypotheses tested in this chapter. A.K, C.E, M.K, C.L.G, and T.H conducted the fish collections, dissections, and parasite screening. C.L.G carried out sex-typing and genotyping. A.K carried out all microbial laboratory work. A.K conducted data analyses and writing. A.K and C.E interpreted the results. Comments and manuscript edits by C.E, L.H and C.H.

Chapter 5: A.K and C.E conceived the experiment. A.K, C.E, Leila Fouda, Thomas Reischig, Albert Taxonera, Silvana Monteiro Roque, Leno Dos Passos, and Lino Duarte collected samples. A.K conducted all laboratory work. Doko-Miles Jackson Thorburn assisted in bioinformatics. A.K conducted data analyses and writing. A.K and C.E interpreted the results. Comments and manuscript edits by C.E, L.H and C.H.

General Conclusion: A.K conducted writing. C.E gave comments and edits on manuscript.

In memory of those we lost along the way,

Leigh Hart

Daphne Hart

Steve Le Comber

"Many of life's failures are people who did not realize how close

they were to success when they gave up."

--Thomas Edison

Abstract

Microbes are ubiquitous, and are often found in close associations with a host, where they affect its physiology, immune functions and even behaviours. In this thesis, I explored host-microbe interactions in wild vertebrate species to disentangle the role of genetic and environmental determinisms of those interactions. In Chapter Two, together with my collaborators, I brought evidence for the evolution of local adaptation of three-spined stickleback (*Gasterosteus aculeatus*) gut microbiomes through population-specific microbiomes, caused by localised environmental pressures and/or genetic determinism. In Chapter Three, I established the baseline relationship between parasites and stickleback diet, as both can impact host-microbe interactions. I confirmed that feeding ecology is not independent of host-parasite interactions and both parasite resistance and feeding ecology evolve under local adaptation. Chapter Four highlighted how the host's microbiome is influenced by genotype-by-environment interactions, with changes in the microbiome correlating with interactions between host evolutionary lineage, local environment, and seasonal variation. Interestingly, microbial diversity decreased with increased parasite infections, suggesting intricate host-parasite-microbe interactions. Additionally, I found evidence that a host's microbiome is linked to feeding ecology, but the direction of this relationship was context-dependent. Finally, in Chapter Five, I tested the generality of the conclusions obtained in the fish system by changing host species to the philopatric and locally-adapted loggerhead sea turtle (*Caretta caretta*) nesting at the Cabo Verde Archipelago. I found population-specific cloacal microbiomes among closely related nesting groups and as well as host-parasite-microbe interactions. Overall, this thesis focused on teasing apart the diverse determinisms of wild host-microbe interactions. It relied on a series of field experiments and sampling of wild individuals and ultimately shows how fundamental the role of the host microbiome is for species evolution.

Acknowledgments

When people say you should pick the supervisor, not the PhD topic, they were completely correct when it comes to my wonderful, long-suffering supervisor Christophe Eizaguirre. None of the achievements from my PhD (or MSc) would have been possible without you. I cannot put into words how grateful I am for all you have done for me. Without you pushing me I wouldn't have started the PhD, let alone made it to the end and for that, I will forever be indebted to you. It's been an honour to work with you, your passion for research is infectious and I am a better scientist for knowing you. I truly hope that you have enjoyed working with me as much as I have enjoyed working with you, ignoring the odd disagreement here and there! Thank you for your endless support and for always pushing me.

To China Hanson, you helped open my eyes to the wonderful world of microbiology and were always there with a reassuring message if I needed it. To Lee Henry, thank you for introducing me to how truly amazing bacteria can be on the first day we met running around East London collecting aphids. Thank you to all those involved with London NERC DTP for seeing potential in me and giving me this amazing opportunity to develop and grow.

I have had the honour to work with some amazing scientists at QMUL. Special thanks need to go to Ian Sanders, Chloe Economou, and Phil Howard. Without you three I would have been lost in the lab, thank you for all the advice, knowledge, and support you have given me over the years.

I have met some truly wonderful people during this journey, many of which be my friends for life. My biggest thanks has to go to Liam Fitzpatrick and Emma Lockley. You two have somehow managed to pull me out of every hole I have gotten myself into, I will never be able to repay you both for your kindness, friendship, and love. Special thanks also have to go to Miles Thorburn for putting up with my ramblings over many beers many beers, I don't like to admit it, but I am rather fond of you. To Leila Fouda, sharing a tiny, very sandy box room with you was a highlight of the PhD, thank you for being a wonderful friend. To my lovely lab group, Perla Roman, Ceci Hijar, Liam Dickson, Giacomo Vitali, and Stuart Negas thank you for making work enjoyable. Thanks to Joe Williamson for keeping me sane throughout the PhD, I can't wait to have lots of whiskey with you again. To Anna Cutmore, Mattie Brindle, Dan Nicholson, Ellen Coombs, Rosie Williams, Claire Routledge, Thomas Baird, Michael Stevens, and the rest of the DTP, I couldn't have asked for a better group of people to walk alongside during this journey. Special thanks to Sally Faulkner and Rosie Drinkwater for the multitude of

pep talks they gave me. The transient nature of a PhD means you are constantly saying goodbye to people you have grown close to, but those bonds are not forgotten, so thanks go to Sahmorie Cameron, Dominik Schmid, Leah Lewington Pearce, Tor Kemp, Esther Odekunle, Curtis Horne, Emeline Favreau, Hanrong Tan, Lowri Evans and everyone else at Queen Mary for welcoming me with open arms.

A huge thanks to my non-academic friends, especially Beth Foley, Ele Mather, Georgia Buckland and the rest of the girls for cheering me on every step of the way. I'm excited to get back to being fun and spontaneous with you again.

Last but most certainly not least, my loving family. To my Mum and Dad, your unwavering belief in me has always pushed me to strive for more and to see myself as you so clearly see me. To my sister, Georgina, and her partner Michael, thank you for always being there when I needed you. I would be lost without all four of you.

Table of Contents

List of Figures

List of Tables

Chapter 1: General Introduction

Since Charles Darwin, the study of species evolution has moved forward, particularly facilitated by the advance in genomics and ecological techniques. Over recent decades, it has become possible to not only determine a population's adaptive potential, but also to investigate some of the most cryptic biotic interactions, such as those between hosts and microbes. In this introduction, I will first explain the evolutionary concept associated with local adaptation, before focusing on the specific role microbes play within a host and what selection pressures influence the diversity and structure of microbiomes.

1.1 Local adaptation

Local adaptation is a key component of adaptive evolution, occurring when a resident population has increased fitness in their local environment compared to migrants (Kawecki and Ebert, 2004, Savolainen *et al.*, 2013, Sobel *et al.*, 2010). Local adaptation is often represented as a genotype-by-environment interaction, whereby the fitness of an individual in a specific location is related to its genotype and surrounding environmental conditions (Des Marais *et al.*, 2013). When abiotic environmental selection pressures such as temperature, salinity, and photoperiod vary across time and space, local conditions will determine what traits will be favoured by natural selection (DeFaveri and Merila, 2014, Griffith and Watson, 2005, Jackson *et al.*, 2020). For example, the fry of three-spined stickleback (*Gasterosteus aculeatus*) from different salinity regimes show local adaptation to their native salinity when placed in low, medium, and high salinity treatments (DeFaveri and Merila, 2014). This research found an increased probability of survival in the high salinity treatment of native fry in comparison to those from low salinity backgrounds. As natural selection acts on specific traits, changes in allele frequencies underlying those traits can shift the population toward a local optimum. Over time, natural selection acting in different locations can lead to adaptive divergence of trait means and allele frequencies (Kawecki and Ebert, 2004).

Local adaptation can emerge from local abiotic selection but also from biotic interactions. The influence of host-parasite interactions on the evolution of local adaptation of a host population has repeatedly been studied (Barber, 2013, Eizaguirre *et al.*, 2012a, Kalbe and Kurtz, 2006, Kaufmann *et al.*, 2017, Lenz *et al.*, 2013). This is because parasite and pathogen communities will also depend on the local ecology (Thomson *et al.*, 2007). How host-parasite interaction is associated with local adaptation has even been experimentally determined in the field associating fitness to diversity of immune genes (Eizaguirre *et al.*, 2012a, Eizaguirre *et al.*, 2012b).

Another biological interaction that has the potential to contribute to the evolution of host local adaptation is that of host-microbe interactions. It has been suggested that without a good understanding of a host's associated microbes, it is impossible to fully understand host evolution (McFall-Ngai *et al.*, 2013, Rosenberg and Zilber-Rosenberg, 2016, Rudman *et al.*, 2019). Changes in microbiomes can have a significant impact on their hosts, altering their evolutionary trajectories (McFall-Ngai *et al.*, 2013). For example, microbiomes are capable of influencing the phenotype of their host but this often depends on the ecological context and their potential to influence host evolution remained relatively unexplored (Koskella *et al*., 2017, McFall-Ngai *et al.*, 2013, Rudman *et al.*, 2019). Theories like the hologenome suggest that both the eukaryotic host and their microbiomes are combined into a singular evolutionary unit suggesting that the evolution of a host and their microbes are strongly interlinked (Bordenstein and Theis, 2015, Theis *et al.*, 2016). Henry *et al*. (2021) summarised the current literature and theories surrounding the coevolution of hosts and their microbiomes suggesting hosts may utilise locally adapted microbes, found in the surrounding environment to shift phenotypic means between different populations potentially leading to local adaption of a host. Selection pressures that are acting upon the host will also be acting upon the bacteria, however bacteria have larger population sizes, shorter generation times and may evolve

adaptive functions faster than host populations (Ferreiro *et al*., 2018, Koonin and Wolf, 2012) so if a host is able to utilise these microbes to their advantage consistently they may evolve mechanisms that could lead to a constant symbiotic relationship (Sachs *et al*., 2011). An example of this is the bean bug *Riptortus pedestris*, which is a common pest of legumes, and is irradicated with pesticides, however they can gain resistance to these by obtaining locally adapted pesticide degrading *Burkholderia* from the surrounding soil (Itoh *et al*., 2018).

A multitude of selection pressures acting upon a host have been shown to be mediated by the presence of locally adapted microbes, for example the presence of heat tolerant *Curvularia* found in geothermal soils increased the thermotolerance of non-adapted tomato plants improving their survival, whilst the introduction of *Curvularia* which was not locally adapted to the increased temperatures had no impact on plant survival (Rodriguez *et al*., 2008). Additionally, a field reciprocal transplant experiment examined the role of soil microbes in the local adaptation of St. John's wort (*Hypericum perforatum*) to the stressful habitat of limestone barrens (Petipas *et al.*, 2020). Local seedlings had increased survival and germination rates when planted with microbes originating from the same environment, demonstrating microbe-mediated local adaptation. Another study examined the kidney microbial communities of sympatric pairs of dwarf and normal lake white fish (*Coregonus clupeaformis*) across five different lakes. They aimed to test whether microbial diversity and composition evolved in parallel within the sympatric pairs across multiple habitats (Sevellec *et al.*, 2014).While there was no clear evidence that the kidney microbiome evolved in parallel across the systems, they found a genotype-by-environment interaction whereby the difference in microbial community between morphs (dwarf or normal) was lake-specific (Sevellec *et al.*, 2014). Similarly, Sullam *et al.* (2015) explored the potential parallel evolution of the gut microbiomes of two recently diverged Trinidadian guppy, *Poecilia reticulata*, ecotypes that exhibit different diets, life history traits and morphologies across low and high-predation habitats. They found

variation in microbial communities over time, across streams and among ecotypes but this was not parallel across all streams (Sullam *et al.*, 2015). These specific examples show that host-microbe interactions can be associated with local adaptation.

There are multiple theories around how a two linages coevolve, this is the reciprocal evolution of a population in response to another and can be applied to host-microbe interactions (Zaneveld *et al.*, 2008).The Red Queen's hypothesis, antagonistic coevolution could potentially occur in the presence of a pathogen where the fitness of one species increases at the cost of the other species (Van Valen, 1974). Alternatively, mutualistic coevolution may occur in host-microbe relationships, where there is obligate symbiosis and metabolic collaboration as both species benefits from the presence of the other (Herre *et al.*, 1999). A well understood example of this is the pea-aphid who has evolved specialised cells called bacteriocytes to hold a vertically transmitted endosymbiont, *Buchnera aphidicola*, who in return synthesize amino acids that the aphids require (Braendle *et al.*, 2003). A third theory is that of the hologenome, where the host and their microbiome can act as single unit of selection, as both the genomes of the host and their microbiome can influence host phenotype (Bordenstein and Theis, 2015, Theis *et al.*, 2016), however these are likely to be spatially and temporally heterogeneous as well as influenced by neutral processes (Koskella and Bergelson, 2020, Theis *et al.*, 2016).

At this stage, before diving deeper into the mechanisms of host-microbe interactions, I want to clearly define what is meant by "microbes".

1.2 Microbes

Microbes is a generalised term for bacteria, fungi, archaea or protists, but the main focus of host-microbe interactions is often bacteria due to their diversity and abundance. Bacteria are omnipresent in the biosphere and make up at least 15% of the biomass on Earth, while animals only constitute ~ 0.36% (Bar-On *et al.*, 2018), with their habitats ranging from the greatest depths of the world's oceans (Parkes *et al.*, 1994) and hydrothermal vents (Taylor *et al.*, 1999), to the upper reaches of the atmosphere (Smith,

2013). Microbes can be free-living or aggregate-attached, with many microbes living on or within other organisms, which are referred to as their host (Gilbert *et al.*, 2012). In general, the microbe community is referred to as a microbiome, i.e. the combined genetic material of microorganisms in a particular environment (Lederberg and McCray, 2001). Most microbes that form a host species' microbiome can have biphasic lifecycles, where part of the microbe's life is spent in a different environment, i.e. as free-living bacteria or associated with a different host (Obeng *et al.*, 2021). The acquisition of microbes by their host can occur through several potential routes. In mammals, this may occur through both vertical transmission, from mother to offspring (Dominguez-Bello *et al.*, 2010, Vaishampayan *et al.*, 2010), and horizontal transmission, through social interactions and shared environments (Moeller *et al.*, 2016a, Perofsky *et al.*, 2017). Organisms without live birth or parental care must acquire their bacteria from the environment. For example, the first microbial colonisers of newly hatched fish larvae are sourced from the surrounding water and their food supply (Hansen and Olafsen, 1999, Korsnes *et al.*, 2006, Reid *et al.*, 2009). Scientific knowledge of the mechanisms by which microbes colonise a host, and the impact of host-microbe interactions on both the host organism and the microbiome, is increasing at a rapid pace. There is, however, still plenty that needs to be understood about these complex relationships.

1.3 Host-microbe interactions

Bacteria within a host's microbiome have high levels of phenotypic plasticity, short generation times, and are capable of horizontal gene transfer, resulting in highly dynamic communities that are heavily influenced by local selection pressures (Walter and Ley, 2011). Plasticity might increase host tolerance to environmental changes and potentially contributes to population-level divergence and local adaptation (Alberdi *et al.*, 2016, King *et al.*, 2016, Kolodny and Schulenburg, 2020).

Only a small group of microbes are classified as pathogenic (Hornef, 2015), the majority of microbes living within a host are thought to be commensal or mutualistic (Alberdi *et al.*, 2016, Koskella *et al.*, 2017, McFall-Ngai *et al.*, 2013). Firstly, they can aid pathogen defence through colonisation resistance (Gerardo and Parker, 2014, Parker *et al.*, 2011). While the underlying mechanisms of this are not clear, theory suggests bacterial species outcompete pathogens for niche space through the secretion of antimicrobial peptides (Kim *et al.*, 2017b, Lawley and Walker, 2013). Secondly, links have been identified between a host's microbiome and immune system function (Kelly and Salinas, 2017, Lee and Hase, 2014). For instance, the presence of *Bacillus* and *Lactobacillus* within the intestines of fish stimulate the expression of inflammatory cytokines (He *et al.*, 2017), increase phagocytic activity (Chen *et al.*, 2019) and increase goblet cell formation, which helps produce a protective mucus layer (Topic Popovic *et al.*, 2017). Additionally, host behaviours can be influenced by the microbiome (Davidson *et al.*, 2020, Vuong *et al.*, 2017), as seen in microbiome-related changes in the odour of red harvester ants, which increased the likelihood of the host to be attacked by the rest of the colony (Dosmann *et al.*, 2016). Lastly, microbial metabolites within the gut can be used as a source of nutrients for the host, transforming indigestible food products, such as cellulose and plant-derived pectin, into useable compounds for both the microbes and the host (Bäckhead *et al.*, 2005, Turnbaugh *et al.*, 2006).

Interestingly, microbes also have the ability to cross the parasite-mutualist continuum, becoming pathogenic in the absence of more virulent pathogens or in stressful environments, or becoming beneficial in the reverse situation (Chamberlain *et al.*, 2014). For example, *Caenorhabditis elegans* can carry the weakly pathogenic bacteria, *Enterococcus faecalis*, and upon potential infection of the more virulent *Staphylococcus aureus, E. faecalis* crosses the parasitism-mutualism continuum to protect their host

(King *et al.*, 2016, Rafaluk-Mohr *et al.*, 2018). As a result of these possible transitions, it is difficult to classify bacteria as purely pathogenic or beneficial.

One of the key questions in the field of host-microbe interactions is the role a host and their environment play in structuring microbial diversity and composition. A greater understanding of this will provide insight into host-microbe interactions in their entirety (Adair and Douglas, 2017, Alberdi *et al.*, 2016, Spor *et al.*, 2011).

1.4 What influences the microbiome?

1.4.1 Host influence

Microbiome structure and composition is heavily influenced by host genotype (Figure 1.1, Smith *et al.*, 2015, Steury *et al.*, 2019, Sullam *et al.*, 2012, Wang *et al.*, 2016a). Genetic control can stem from host immunity. For instance, differences within the major histocompatibility complex (MHC) gene region, which only exists in jawed vertebrates, are correlated with the composition of the human microbiome, resulting in the microbiome of genetically similar hosts being more alike than unrelated hosts (Bonder *et al.*, 2016, Chen *et al.*, 2018, Steury *et al.*, 2019). Immunity-based selection by the host favours genotype-specific microbial communities that are stable with a high level of functional redundancy (Ley *et al.*, 2006). Interestingly, the host genotype can have such a strong influence on their microbiome, that even after two decades in shared standardised laboratory conditions on the same diet, two *Hydra* species retained significantly different microbial communities, similar to their wild counterparts (Fraune and Bosch, 2007). Studies on the mammalian microbiomes suggest that host genotypes may be a primary cause of intraspecific gut microbiome variation (Zoetendal *et al.*, 2001; Hildebrand *et al.*, 2013; Linnenbrink *et al.*, 2012). The influence of host genetics, however, appears to be weaker in birds, as bird gut microbiomes are more heavily influenced by their host's diet and geography (Hird *et al.*, 2014, Waite and Taylor, 2014, Waite and Taylor, 2015). This example shows that the strength of genetic influence varies across taxa. Additionally,

sexual dimorphism in the gut microbiome has been shown in largemouth bronze gudgeon and mice (Figure 1.1, Elderman *et al.*, 2018a, Li *et al.*, 2016). Studies in zebrafish, however, found no difference in the microbiome linked to fish sex (Liu *et al.*, 2016b, Stephens *et al.*, 2016). This finding may be linked to the lack of heteromorphic sex chromosomes or single sex-determining locus in zebrafish, suggesting that instead, the genes contributing to the determination of sex are distributed throughout their genome (Liew *et al.*, 2012, Traut and Winking, 2001).

The ecosystem on a leash theory highlights of four additional methods of host genetic control, the evidence for which is summarised in depth by Foster *et al.* (2017). Firstly, a host is able to control immigration of bacteria, this can be through learned behavioural mechanisms, such as avoiding unclean areas or rancid food sources (Welzl *et al.*, 2001) or biological as stomach acid is capable of destroying ingested microbes (Imhann *et al.*, 2016). Secondly, potentially beneficial species can be targeted, through the host providing nutrients to the desired species (Sonnenburg *et al.*, 2005). Hosts can also monitor the bacteria residing within them, mammals are able to identify the location of microbes through their toll-like receptors or monitor the benefits given by certain bacteria (Kiers *et al*., 2003, Vaishnava *et al*., 2011). A clear example of benefit monitoring is that of legumes and bacteria, such as *Bradyrhizobium japonicum* that reside within their root nodules and are provided nutrients by the plant in exchange for fixing nitrogen. If the amount of nitrogen produced by a group of bacteria within a nodule decreases, then the plant will stop the nutrient supply (Kiers *et al*., 2003). Additionally, legumes exhibit another form of control, compartmentalising, where bacteria are contain in specific areas, in this case the plants root nodules (Kiers *et al*., 2003)

Other non-genetic host influences include pregnancy/gestation (Figure 1.1), which has been linked to changes in the diversity and composition of gut microbes in bats (Phillips *et al.*, 2012), mice (Elderman *et al.*, 2018b), and the oviparous eastern fence

lizard (Trevelline *et al.*, 2019). The direction of this change varies across different taxa, with bats showing increased diversity during gestation, whilst the microbial diversity of mice and lizards decreased (Elderman *et al.*, 2018b, Phillips *et al.*, 2012, Trevelline *et al.*, 2019). Interestingly, the changes in microbial diversity observed during pregnancy and lactation in Phayre's leaf monkeys (*Trachypithecus phayrei*) was negatively correlated with progesterone concentrations, suggesting reproductive hormones may moderate microbial variation (Mallott *et al.*, 2020). Additionally, microbial diversity and composition vary with the developmental stage of a host species (Avershina *et al.*, 2016, Kohl *et al.*, 2013, Yan *et al.*, 2016). Such a pattern was identified across three freshwater fish, *Ctenopharyngodon idellus*, *Siniperca chuatsi*, and *Silurus meridional*, where microbial community composition differed between all three developmental stages (larvae, juvenile and adult), despite similar rearing environments (Yan *et al.*, 2016). Furthermore, bacterial diversity decreased as all three fish species developed although the underlying mechanism has still not been elucidated to date (Yan *et al.*, 2016). Variation driven by non-genetic host effect must therefore be considered as a possible confounding factor when examining host-microbe interactions if the research question does not specifically aim to explore the impact of non-genetic effects.

Microbiome diversity and structure can change with host behaviour and social ranking (Figure 1.1). For example, in male cichlid fish (*Astatotilapia burtoni*) which have strong social ranking, subordinate individuals harboured lower microbial diversity and more pathogenic clades than dominant males, potentially reducing fitness of lower-ranking males (Singh *et al.*, 2019). Increased social contact in chimpanzees, calculated using the proportion of time individuals spent together, was positively correlated with species richness within the gut microbiome resulting in homogeneity in microbial communities among individuals (Moeller *et al.*, 2016b). Consequently, laboratory studies which restrict normal social structure and behaviours through housing individuals

separately, or contrastingly, housing study organisms in close contact to each other, may be impacting the host microbiome unintentionally.
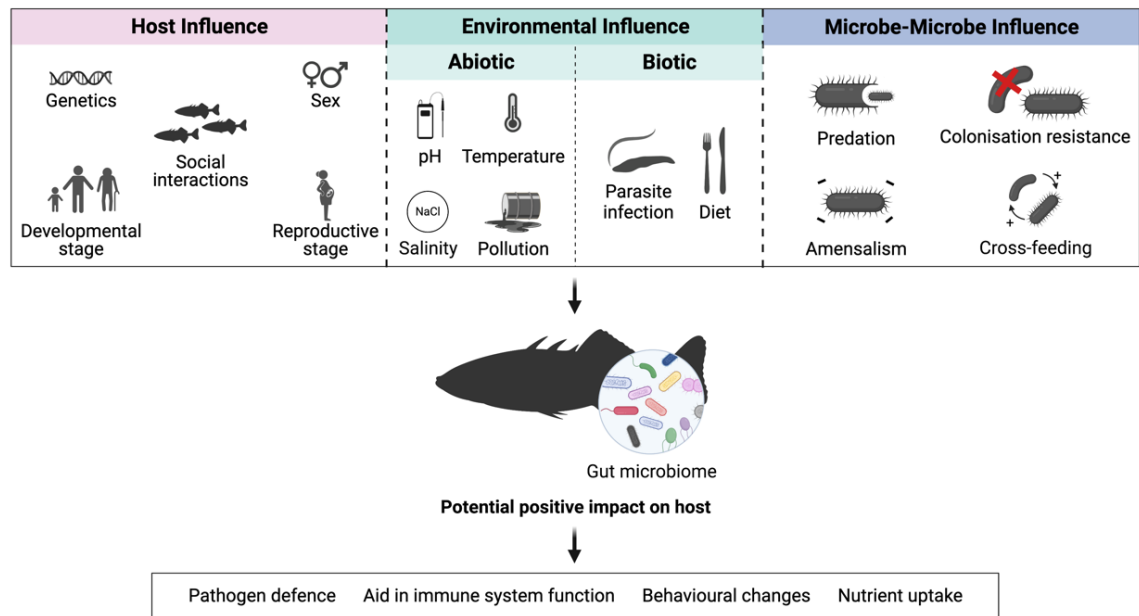
Figure 1.1: A schematic diagram of the influential factors acting upon the host gut microbiome and examples of potential positive impact the gut microbiome may have on its host. Created in biorender.com.

1.4.2 Host environment

1.4.2.1 Abiotic influence

The environment, including the abiotic and biotic factors in which a host resides not only influences the host directly, but also impacts its microbiome (Figure 1.1, Spor *et al.*, 2011). In particular, the addition of pollutants, plastics, heavy metals, and pesticides to an environment lead to variation in a host's microbiome; exposure to these chemicals correlated with decreases in the abundance of certain bacteria whilst others increased (Jin *et al.*, 2018, Kan *et al.*, 2015, Meng *et al.*, 2018). Similarly, antibiotics can cause dysbiosis in the gut microbiome of zebrafish with this change associated with a reduction in nutrient absorption (Zhou *et al.*, 2018). Laboratory mice given acidic drinking water showed a decrease in microbial diversity in comparison to those given neutral pH water (Sofi *et al.*, 2014). Sullam *et al.* (2012) showed salinity can influence the microbiome of fish regardless of its phylogeny as freshwater fish harbour more similar microbiomes than

those in marine habitats. Additionally, seasonal changes can influence the host microbiome through temperature changes but also variation in food consumption and feeding strategy (Al-Harbi and Naim Uddin, 2004). For example, the abundance of bacteria in the intestine of tilapias was lower in winter compared to other seasons (Al-Harbi and Naim Uddin, 2004). Increased temperatures have been linked with reduced microbial diversity in mice (Chevalier *et al.*, 2015), chickens (Zhu *et al.*, 2019), lizards (Bestion *et al.*, 2017), and salamanders (Fontaine *et al.*, 2018). Contrarily, temperature had no significant effect on the microbial diversity of tadpoles or cows, suggesting the influence of temperature may be species-specific (Kohl and Yahn, 2016, Tajima *et al.*, 2007). Consequently, care must be taken to maintain standard abiotic conditions when carrying out laboratory experiments. Within wild field-based experiments, if microbial data is to be collected in a time-series care must be taken to either sample at the same time period or carry out pilot studies to identify if there are temporal effects influencing the microbiome.

1.4.2.2 Captive vs wild host microbiomes

Because of the importance of abiotic factors, it is not surprising that the rearing environment of a host alters its host-microbe interactions. As such captivity can influence the diversity and community structure of the microbiome, as observed in mammals (Clayton *et al.*, 2016, Gibson *et al.*, 2019, McKenzie *et al.*, 2017), fish (Eichmiller *et al.*, 2016, Restivo *et al.*, 2021), reptiles (Keenan *et al.*, 2013), and birds (Oliveira *et al.*, 2020, Wang *et al.*, 2016b). Animals in captive environments experience changes in diet, antibiotic exposure, increased stress, human contact, and reduced habitat variation, all of which have the potential to alter the structure of a host's microbiome (McKenzie *et al.*, 2017, Portz *et al.*, 2006). Reduced microbial diversity has been observed across a range of captive canids, primates, and equids when compared to their wild counterparts (McKenzie *et al.*, 2017). Not all species, however, show changes between wild and

captive hosts; the same study showed that the microbial diversity of bovids, giraffes, anteaters, and aardvarks remained similar to their wild counterparts, regardless of captivity status (McKenzie *et al.*, 2017). Additionally, Ley *et al.* (2008a) did not find any association between animal captivity and the faecal microbiome of 59 mammal species. These mixed results highlight that care must be taken when sampling the microbiome of captive organisms from zoos or within laboratories and indeed emphasises that the study of wild microbiomes is crucial for drawing ecologically and evolutionarily relevant inferences (Hird, 2017). The importance of studying the wild microbiome has been recognised in recent years with an increase in studies focusing on both classic model organisms and more novel organisms. Maurice *et al*. (2015) found that seasonal changes was the dominant factor influencing the gut microbiome of wild wood mice (*Apodemus sylvaticus*) over a two-year period, suggesting a shift in diet may be a reason behind this, such information would have been overlooked in captive situations. Frog species show a similar seasonal changes in their skin, stomach and gut microbiome, however this is likely driven by behavioural changes in the habitat in which they reside, moving from ponds in the spring and summer into drier conditions for autumn ready to hibernate (Xu *et al.*, 2020). Additionally, a study focused on the gut microbiome of small wild mammal species across multiple habitats, found species-specific microbiomes even within a complexed shared habitat (Knowles *et al.*, 2019).

1.4.2.3 Diet

Host diet has been identified as one of the most influential factors associated with changes in the gut microbiome (Figure 1.1, Foster *et al.*, 2017, Li *et al.*, 2017, Muegge *et al.*, 2011, Sullam *et al.*, 2012), and has been identified as one of the main factors influencing the disparity between captive and wild microbiomes (McKenzie *et al.*, 2017). Food sources act as a source of colonising bacteria (Costello *et al.*, 2012) containing different nutrients to be assimilated by different bacterial species within the host, which

logically alter the composition of the gut microbiome (David *et al.*, 2014). An example of the strong influence diet has on the host microbiome is observed in both mammals and fish, whereby herbivores often have higher microbial diversity than omnivores or carnivores (David *et al.*, 2014, Larsen *et al.*, 2014, Li *et al.*, 2014, Muegge *et al.*, 2011). Particularly, there is a general increase in abundance of cellulose-degrading bacteria *(Clostridium, Citrobacter* and *Leptotrichia)* in herbivorous fish, whilst *Cetobacterium* and the protease-producing *Halomonas* dominate carnivore microbiomes, likely assisting in nutrient uptake (Liu *et al.*, 2016). Diet has heavily contributed to the evolution of the gut microbiome of myrmecophagous (ant and termite-eating) mammals, with the convergence of microbial communities across species and global distribution (Delsuc *et al.*, 2014). Such convergence could suggest the use of symbiotic bacteria to aid in digestion of exoskeletons. These results show diet adaptation can be a major factor of gut microbiome composition over evolutionary timescales. Many host-microbe-diet studies have focused on the broad influence of dietary groups on the host microbiome (i.e. plant-based diets vs animal-based diets, or high fat vs high fibre diets) and are often carried out in controlled conditions (David *et al.*, 2014, Heinritz *et al.*, 2016). This is not directly representative of the microbe-diet interactions within natural systems, whereby organisms consume a wide range of food sources, thus further testing is required.

Host genetics can play a role in the evolution of feeding strategy, revealing the influence of diet on the microbiome is genotype-by-environment dependent. For instance, sympatric benthic-limnetic three-spined stickleback show genetic and morphological differences, as well as different feeding strategies correlating with differences between their gut microbiomes (Rennison *et al.*, 2019a, Schluter, 1995). Sex-specific diets can also lead to sex-specific microbiomes, another form of genotype-by-environment interaction involving host diet (Bolnick *et al.*, 2014c). As diet has the potential to strongly influence a host's gut microbiome, the need to quantify the feeding ecology of a wild

organism is paramount. This can be achieved in several ways, through feeding observations, gut content analysis or stable isotope analysis (Rudnick and Resh, 2005). Stable isotope analysis is an ideal way to quantify host diet and can also be linked to parasite infection status of an individual (Box 1).

Box 1: Stable isotope analysis

Stable isotope analysis (SIA) allows for long-term assessment of an individual's diet, and is predicated using the concept that a consumer's cells are synthesised from materials assimilated from their diet (Lorrain *et al.*, 2002, Post, 2002). Most elements in the periodic table have several naturally occurring stable isotopes. The two most commonly used for assessing feeding ecology of an organism are carbon (C) and nitrogen (N) (Crawford *et al.*, 2008).

Mass spectrometry is used to measure stable isotope ratios by separating the different isotopes based on their mass-to-charge ratio. Due to isotopic variation between prey items, and the fractionation that occurs following key biological processes such as assimilation, the proportion of littoral carbon in a host's diet as well as their trophic position can be estimated (Figure 1.2, Matthews *et al.*, 2010, Post, 2002). The $^{13}C/^{12}C$ ratio ($\delta^{13}C$) describes the organic carbon source of a food web, with the $\delta^{13}C$ value of an organism regularly found to be within 1 ‰ of the value of its food source (DeNiro and Epstein, 1978). The proportion of littoral carbon within an individual's diet can be calculated by comparing a consumer's $\delta^{13}C$ with $\delta^{13}C$ of the primary consumers within a system (Matthews *et al.*, 2010, Post, 2002). Nitrogen ($^{15}N/^{14}N$) isotope ratios ($\delta^{15}N$) are enriched by ~3.4 ‰ in consumers in comparison to their food source, allowing for individual trophic position to be calculated (Matthews *et al.*, 2010, Post, 2002).

Results from stable isotope analysis can be correlated to the microbial diversity of a host and can then be tested for the link between feeding ecology and microbial diversity (Bolnick *et al.*, 2014b, Bolnick *et al.*, 2014c, Gongora *et al.,* 2021).
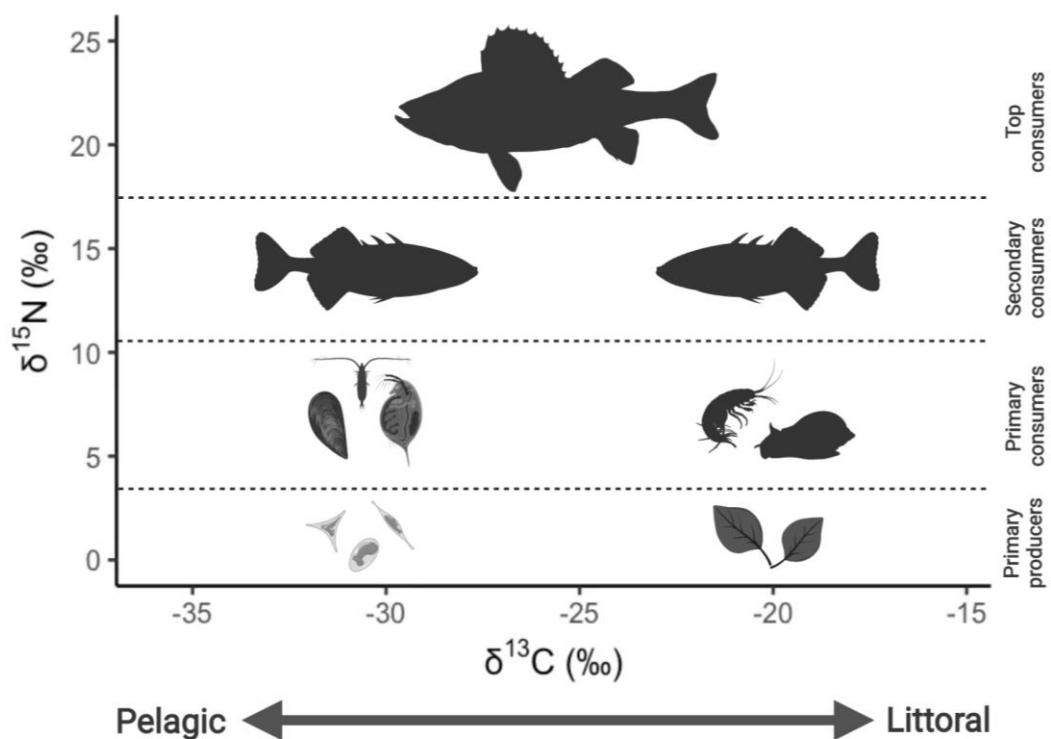
Figure 1.2: A schematic stable isotope biplot showing the distinction between pelagic and littoral food-webs in aquatic systems by carbon ($\delta^{13}C$) and nitrogen ($\delta^{13}C$) isotope ratios. Created in biorender.com.

1.4.2.4 Parasites

Parasites are one particular element of the environment that correlates with feeding ecology and impacts host-microbe interactions. Similar to microbes, parasites are ubiquitous, and their abundance and community structure depend on ecological conditions (Poulin, 2011). Some parasites are trophically transmitted, while others will directly infect their host (Barber, 2013, Stewart *et al.*, 2017). Given the costs of parasitism for their host, they form a significant selection pressure that must be considered as an integral part of the environment which hosts are exposed to (Barber, 2013, Eizaguirre *et al.*, 2012b, Milinski and Bakker, 1990). Importantly for one of the themes of this thesis, host-parasite interactions have frequently been linked to local adaptation of host populations (Eizaguirre *et al.*, 2012a, Kalbe and Kurtz, 2006, Kaufmann *et al.*, 2017, Lenz *et al.*, 2013, Summers *et al.*, 2003). However, it is only relatively recently that the importance of the relationship between a host, their parasites and their microbiome has been brought to light (Figure 1.1, Dheilly, 2014). Findings on the influence of parasite infection on host microbial diversity differ greatly, with studies showing both increased (Lee *et al.*, 2014, Rosa *et al.*, 2018) and decreased microbial diversity (Houlden *et al.*, 2015, McKenney *et al.*, 2015) - suggesting these relationships could be specific to the organisms involved in the interaction.

In general, it seems that parasites can directly modify the host microbiome by disrupting its diversity and community to create more beneficial conditions for its establishment (Dheilly *et al.*, 2015). An interesting example is that of the microsporidian parasite, *Paranosema locustae*, which infects locusts and can modify the host's hindgut microbiome and host behaviour (Shi *et al.*, 2014). Specifically, *P. locustae* acidifies the hindgut and modulates the host's immune response, reducing the abundance and growth of microbes and allowing higher infection rates. In turn, the acidification of the hindgut

prevents locusts from swarming, as their microbiome is involved in producing pheromones that trigger swarming behaviour (Shi *et al.*, 2014).

Parasites can also use their own microbiota as a biological weapon (Dheilly *et al.*, 2015). Entomopathogenic nematodes have a mutualistic relationship with enterobacteria *Photorhabdus* and *Xenorhabdus* (Adams *et al.*, 2006, Boemare and Akhurst, 2006). Upon entering the host, *Photorhabdus* and *Xenorhabdus* are released from the nematodes' intestine, creating toxins that kill the host and help with tissue degradation (Boemare and Akhurst, 2006). This shows mutualistic coevolution between nematodes and microbes, where the nematode provides a home in which the bacteria reside, and in turn the bacteria provide a service. The numerous ways a host's microbiome can both, directly and indirectly, assist in a pathogen's ability to infect their shared host have recently been reviewed by Stevens *et al.* (2021).

Microbes can also aid in parasite defence, resulting in coevolution between the host and the microbe, as well as between the microbe and the parasite (reviewed in Ford and King, 2016). Defensive microbes can protect their host from parasites directly through (i) hyperparasitism, where microbes parasitise the parasites, reducing the parasites survival (Davies, 2009, Tollenaere *et al.*, 2014), (ii) outcompeting parasites for a specific niche, through colonisation resistance (Gerardo and Parker, 2014, Parker *et al.*, 2011), and (iii) by producing antibiotics or bacteriocins which can reduce parasite growth rate or kill them (Gerardo and Parker, 2014, Mideo, 2009). For instance, the uropygial gland secretions of hoopoes, *Upupa epops,* contain *Enterococcus faecalis,* which produces bacteriocins, such as enterocin MR10, that protect the bird's feathers from a wide range of microbial parasites, such the keratinolytic bacterium *Bacillus licheniformis* which can degrade the keratin in feathers (Ruiz-Rodriguez *et al.*, 2013, Ruiz-Rodriguez *et al.*, 2009). Microbes can evolve rapidly to both their host and their parasites because of the microbes' short-generation times, high levels of phenotypic plasticity and large

population sizes, which could aid in host adaptation (King *et al.*, 2016). Therefore, understanding the host-parasite-microbe interactions of wild organisms is important, as the complexity of natural systems is difficult to emulate *in vitro,* and so the influence of a diverse array of co-infecting parasites could be missed in studies where a singular parasite is used.

Noteworthy, the relationship between host diet and parasite infections can be explored using stable isotope data (Lockley *et al.*, 2020). Exploring the relationship between host-microbe interactions, host diet and parasite infections is a key question in the field of host-microbe interactions.

1.4.3 A combined influence

The relative contribution of each of the determining factors has on the host microbiome remains to be fully elucidated. For example, recent studies of the human gut microbiome suggest host environment is the stronger influencing factor, whilst other studies identify host genetics as more important, showing inconsistency in the conclusions drawn (Garud and Pollard, 2020, Rothschild *et al.*, 2018). Additionally, genotype-by-environment interactions have been shown to influence the host microbiome, making it even harder to disentangle which factor is most important in determining the microbiome structure (Gallart *et al.*, 2018, Glasl *et al.*, 2019). A recent genome-wide association study estimated that approximately 10-20% of the variation in the human gut microbiome is driven by environmental factors, whilst host genetics only accounts for 10% (Wang *et al.*, 2016a).

The combination of abiotic and biotic selection pressures can result in the local adaptation of host populations, but what would we expect to observe if local adaptation is acting upon the microbiome? It is anticipated that if the evolution of a host and their microbiome is linked, then consistent differences within in the host microbiome would be observed. So within a locally-adapted population we would expect to see population-specific microbiomes, with individuals from the same population having more similar

microbiomes than to different populations (Delsuc *et al.*, 2014, Kohl *et al.*, 2018, Sevellec *et al.*, 2014).

1.4.4 Microbe-microbe interactions

Microbe-microbe interactions are outside of the scope of this thesis; however, they are known to impact host microbiome diversity and composition in some circumstances, and therefore they require a brief explanation. Microbe-microbe interactions occur when microbes interact with other microorganisms within their immediate environment (Figure 1.1, Barton and Northup, 2011). The nature of the interaction, positive or negative, will depend on the types of microorganism present, as well as their abundance (Barton and Northup, 2011). Cross feeding is a positive interaction where metabolites created by one bacteria can be utilised by another, as demonstrated by *Bacteroides thetaiotaomicron* and *Faecalibacterium prausnitzii* in rats, where *B. thetaiotaomicron* produces acetate and *F. prausnitzii* consumes it (Wrzosek *et al.*, 2013). Some microbes are predatory, for example, *Myxococcus xanthus* consumes other microbial cells within its immediate environment (Berleman and Kirby, 2009, Keane and Berleman, 2016). Amensalism is another negative interaction, where bacteria produce metabolites to suppress others, for instance, *Clostridium scindens* suppresses *Clostridium difficile*'s growth using bile-derived metabolites within the gut of humans and mice (Buffie *et al.*, 2015). Microbes can perform colonisation resistance and defend a host from potential pathogens. Whilst the underlying mechanisms remain unknown, it is thought they utilise a combination of nutrient competition and secreting antimicrobials (Ducarmon *et al.*, 2019).

1.5 Sampling the microbiome

1.5.1 Culture-dependent vs culture-independent methods

There are several technical approaches for determining the microbial community of an organism. These fall into two broad categories, known as culture-dependent and

culture-independent methods (Aagaard and Segars, 2014, Isaacson and Kim, 2012). Culture-dependent studies extract and cultivate bacteria on a range of different growth media. One of the major drawbacks of this approach is a tendency to detect only a selection of microbes (Austin, 2006; Wu *et al.*, 2010; Larsen *et al.*, 2013): only 0.1 - 10 % of microbes seem viable for culture, due to a lack of understanding or ability to replicate specific growth conditions *in vitro* (Amann *et al.*, 1995, Stewart, 2012).

As such, culture-independent methods have grown in popularity, as they appear less limiting in their capacity to detect diversity. Methods include fluorescence in situ hybridization (FISH), denaturing gradient gel electrophoresis (DGGE), temporal temperature gradient electrophoresis (TTGE), metagenomics and targeted amplicon sequencing (Giraffa and Neviani, 2001, Zhou *et al.*, 2014). In recent years, advances in next generation sequencing (NGS) have allowed for fast, accurate and reasonably low cost sequencing of entire microbial communities using targeted amplicon regions, making it one of the most widely used methods today (Barko *et al.*, 2018, Bik, 2016, Foster and Bell, 2012, Ghanbari *et al.*, 2015).

1.5.2 16S rRNA

The small subunit ribosomal RNA (16S rRNA) gene is often the target of NGS when identifying microbial abundances and communities (Figure 1.3). This gene is ubiquitous across all bacteria (Woese *et al.*, 1990), is approximately 1600 base pairs long and consists of nine hypervariable regions and ten conserved regions (Kim *et al.*, 2011). This genomic architecture allows for the identification of bacteria at species level by comparing sequences to large reference databases of bacterial taxonomy, but still allows the use of a universal primer to guarantee broad amplification and comparable results (Stackebrandt and Goebel, 1994; Coenye and Vandamme, 2003). The V1 to V4 regions provide the most accurate estimates of bacteria abundance and diversity in a sample and therefore these regions are recommended for studying host-microbe interactions (Kim *et al.*, 2011). The constant rate of near neutral evolution in the 16S rRNA gene further

permits inferences of phylogenetic relationships between bacteria taxa and is a powerful tool that can provide a comprehensive view of all microbial residents that live within hosts (Woese *et al.*, 1990). Consortiums like the Earth Microbiome Project, the Human Microbiome Project and the Parasite Microbiome Project are working to increase the amount of host microbiome data, as well as to the improve the field through the standardisation of procedures and analysis such as a recommended extraction and amplification protocols, however in such a rapidly progressing field this remains a challenge (Dheilly *et al.*, 2019, Gevers *et al.*, 2012, Gilbert *et al.*, 2014, Gilbert *et al.*, 2018, Schloss *et al.*, 2011, Turnbaugh *et al.*, 2007).
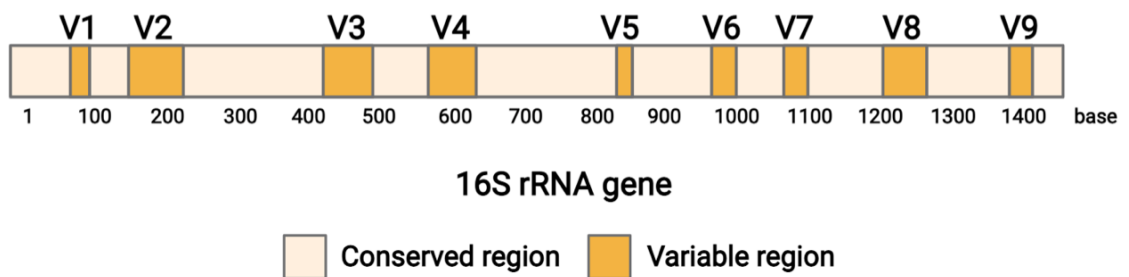
Figure 1.3: A schematic representing the 16S rRNA gene, conserved regions (light orange) and variable regions V1-V9 (dark orange) are displayed. Redrawn in part from Fukuda et al. (2016). Created in biorender.com.

1.6 Analysis of microbial data

1.6.1 Bioinformatic tools

As the amount of data generated through next generation sequencing increased, bioinformatic tools that are capable of turning raw sequencing data into biologically meaningful information are needed. This field is ever-growing but currently, the three most commonly used tools are Quantitative Insights Into Microbial Ecology (QIIME1) (Caporaso *et al.*, 2010), QIIME2 (Bolyen *et al.*, 2019) and Mothur (Schloss *et al.*, 2009). These tools can demultiplex, quality filter, classify operational taxonomic units (OTUs) or amplicon sequence variants (ASVs), assign taxonomic and reconstruct phylogenies. QIIME allows for the visualisation and statistical analysis both within and outside of the

quality control pipeline, whereas Mothur creates files that can be used in programmes such as R (Bolyen *et al.*, 2019). QIIME1 and Mothur arbitrarily cluster sequencing reads into clusters of 97% similarity called operational taxonomic units (OTUs). The use of OTUs in microbial research is becoming less common as clustering sequences can mask true biological variation (Edgar, 2017).

OTUs are instead beginning to be replaced by ASVs. ASVs are exact sequences that capture true biological variation present in the data. Unlike OTUs, ASVs are comparable across studies as they are not sample dependent and have a range of additional benefits (Table 1.1, Callahan *et al.*, 2017). One caveat of ASVs, however, is that a bacterial genome can contain multiple ASVs if there are multiple copies of the targeted genetic locus, so it is possible to have more than one ASV for a bacterial species (Callahan *et al.*, 2017). QIIME2, the successor of QIIME1, allows for ASVs to be identified and de-noised using integrated tools such as Divisive Amplicon Denoising Algorithm (DADA2) (Callahan *et al.*, 2016), or Deblur (Amir *et al.*, 2017).

Table 1.1: Differences between operational taxonomic units (OTU) and amplicon sequence variants (ASV). Information summarised from Callahan et al. (2017).

Operational Taxonomic Unit (OTU)	Amplicon Sequence Variant (ASV)
Sequences clustered into similar groups, usually 97% similarity threshold	Exact sequences
Requires a reference database to cluster OTUs	Does not require a reference database to assign ASVs
Difficulty assigning OTUs from understudied environments, due to limited representation in reference databases	ASVs assigned accurately even in understudied environments
OTUs do not capture all biological variation present in the data due to reference bias	ASVs capture all biological variation present in the data
Not fully comparable across studies depending on type of OTU clustering i.e. closed-reference/de novo OTUs	Comparable across studies
Multiple species can be included within an OTU if sequences are similar, underestimating bacterial diversity	The same species can be assigned multiple ASVs if there are multiple copies of the targeted genetic locus, overestimating bacterial diversity
Chimera detection is complex	Chimera detection is simple

1.6.2 To rarefy or not to rarefy?

An intensely debated topic in microbiome research is whether or not to rarefy sequences. Sample library sizes, the number of reads per sample, can vary by orders of magnitude within a single sequencing run (McKnight *et al.*, 2019). To address variable library sizes, rarefaction can be used. This is when samples are adjusted for differences in library sizes to avoid biases (in terms of abundance and diversity) due to variation in library size across samples. During the process of rarefaction, a minimum library size is chosen, and samples that fall below this threshold are removed, while the remaining samples are subsampled to standardise the number of reads across all samples (Willis, 2019). McMurdie and Holmes (2014) suggested that rarefying removes available valid data from the analysis and is therefore inadmissible. They suggested instead using negative-binomial (NB) based methods, such as edgeR (Robinson *et al.*, 2010) or DESeq (Anders and Huber, 2010). However, these methods do not result in a uniform number of reads, which can result in errors in both alpha and beta data analysis (McKnight *et al.*, 2019) and may not be sufficiently robust (Mandal *et al.*, 2015). There is currently no definite consensus within the field, so our research utilises both rarefied and non-rarefied datasets.

1.7 Study Organisms

Research on the structure and function of host microbiomes is heavily biased towards mammals, which make up less than 10% of vertebrates, leaving the other 90% relatively unexplored (Sullam *et al.*, 2012). In addition, most research has been conducted under laboratory conditions. As i) it is impossible to replicate the complexity of natural systems (Hird, 2017) and ii) captivity can alter microbial structure and diversity within a host, *in vitro/ex-situ* studies have limited potential for extrapolation, and instead more attention must be directed towards understanding microbiomes in wild populations (McKenzie *et al.*, 2017). To tackle this knowledge gap, we utilised two model organisms to explore the wild host microbiome, the three-spined stickleback, and the loggerhead sea

turtle. In order to explore the coevolution and possible local adaptation of host-microbe interactions, it is important to focus on locally-adapted study species. Both the three-spined stickleback and the loggerhead sea turtle are locally-adapted to a habitat type or nesting island (Baltazar-Soares *et al.*, 2020, Cameron *et al.*, 2019, DeFaveri and Merila, 2014, Hendry *et al.*, 2002, Lockley *et al.*, 2020, Stiebens *et al.*, 2013b). As with every study organism they have advantages and disadvantages for studying host-microbiome interactions, these are summarised in Table 1.2.

Table 1.2: Table summarising the advantages and disadvantages of using three-spined stickleback and loggerhead sea turtles as study organisms for host-microbiome interactions.

Three-spined Stickleback		Loggerhead Turtle	
Advantages	*Disadvantages*	*Advantages*	*Disadvantages*
Classic model organism so there is a vast amount of literature on their evolution, physiology, and behaviour	Uncertainty on how representative they are when trying to generalise patterns in the vertebrate microbiome	Listed as 'vulnerable' so turtles are an important organism in conservation, studying their microbiome may assist in conservation efforts	Predominantly females sampled due to accessibility, potentially leading to sex bias
Species widely distributed allowing for spatial studies	Fish must be sacrificed to sample gut microbiome	Non-destructive sampling	Contamination risk when collecting samples in the field
Phenotypically distinct ecotypes have different feeding strategies, allowing for the influence of diet to be measured		Exhibit philopatric behaviour so females return to natal beach to lay eggs allowing for sampling on land	
Parasitised by a range of parasites, allowing for host-parasite-microbiota interactions to be measured		Philopatric behaviour reduces gene flow among nesting aggregations	
Can be reared in laboratories for controlled laboratory studies		Nesting groups show local adaption so microbial differences between groups can be measured	
		Most studies of turtle microbiome focus on turtles within captivity so there is plenty to learn about the wild microbiome	

1.7.1 The three-spined stickleback

Figure 1.4: An image of a subadult three-spined stickleback, Gasterosteus aculeatus. Image taken by Pascal Hablutzel.

The three-spined stickleback, *Gasterosteus aculeatus* (Figure 1.4) is a prominent model organism that has been used to explore the ecology, development and evolution of vertebrates (Gibson, 2005, Peichel *et al.*, 2001, Reid *et al.*, 2021). They have a small body size, wide geographic distribution, and can be relatively easily reared in the lab which makes them ideal for both natural and laboratory-based studies (Bell and Foster, 1994). This vast knowledge makes the three-spined stickleback an ideal system to study host-microbe interactions in the wild.

The three-spined stickleback has undergone multiple colonisations of freshwater habitats from marine environments since the last glaciation, resulting in parallel systems across the northern hemisphere of marine and freshwater ecotypes. In the freshwater habitat, further, more recent differentiation occurred and resulted in the parallel evolution of lake and river ecotypes (Bell and Foster, 1994, Eizaguirre *et al.*, 2011, Feulner *et al.*, 2015, Marques *et al.*, 2016). Adaptive standing genetic variation in sticklebacks is observed globally, however fish from the northern Pacific have five times more divergent standing genetic variant loci than fish from the Atlantic basin, with a large number of alleles being lost as sticklebacks expanded out of the Pacific (Fang *et al.*, 2020). Phenotypically distinct ecotype pairs have evolved between specific habitats, including

oceanic–freshwater, lake–river, and benthic–limnetic pairs (Bell and Foster, 1994, Eizaguirre *et al.*, 2011, Feulner *et al.*, 2015, Marques *et al.*, 2016, McKinnon and Rundle, 2002). Ecotypes differ in genetics, morphology, physiology, and behaviour (Bolnick *et al.*, 2018, Cano *et al.*, 2006, Eizaguirre *et al.*, 2011, Hanson *et al.*, 2017, Ravinet *et al.*, 2013, Rennison *et al.*, 2019b). In my thesis we specifically utilise lake-river ecotype pairs to explore our research questions, and so I will focus on these from now on.

Lake-river ecotype pairs have now been described across the entire range of the species' distribution (Berner *et al.*, 2010, Reusch *et al.*, 2001). In general, these ecotypes have different feeding strategies: lake fish feed on both benthic and limnetic food sources, whilst river sticklebacks are restricted to benthic diets (Berner *et al.*, 2008, 2009; Snowberg and Bolnick, 2012). As diet has a strong influence on the gut microbiome (Bloodgood *et al.*, 2020, Muegge *et al.*, 2011, Turnbaugh *et al.*, 2009), this naturally occurring variation makes sticklebacks a good system for exploring host-microbe-diet interactions. In particular, increasing diversity in the stickleback diet has been associated with decreased microbial diversity within the gut (Bolnick *et al.*, 2014b). Noteworthy, the potential influence of the host immune system on the stickleback microbiome has been associated with diet-specific changes which have resulted in different individual immune responses to bacteria (Friberg *et al.*, 2019).

Like for all species, stickleback diet variation is associated with variation in both parasite exposure to trophically transmitted parasites, and feeding near parasitised individuals, which ultimately, we speculate may correlate with differences in microbial composition (Brunner *et al.*, 2017, Locke *et al.*, 2014, Sanchez-Gonzales *et al.*, 2001, Stutz *et al.*, 2014). However, parasite infections can also result in changes in prey consumption, feeding strategy and behaviour (Brunner *et al.*, 2017, Lefevre *et al.*, 2009). Therefore, host diet is affected by, but also affects, parasite load, which suggests that the

complex relationship between a host, their microbiome, diet and parasites must be examined holistically.

Ecotypes pairs experience variation in parasite infections, with generally lake fish being exposed to a greater diversity of parasites than river fish (Bolnick *et al.*, 2020, Eizaguirre *et al.*, 2011, Kalbe *et al.*, 2002, Reusch *et al.*, 2001). These host-parasite interactions have resulted in ecotype-specific parasite resistance that has been associated with immune gene diversity (Eizaguirre *et al.*, 2012b) and expression (Lenz *et al.*, 2013). For instance, higher levels of parasite exposure have resulted in the evolution of increased resistance of lake stickleback against a diverse range of parasites and higher loads than river populations (Eizaguirre *et al.*, 2012a, Eizaguirre *et al.*, 2012b, Lenz *et al.*, 2013, Wegner *et al.*, 2003). Host-parasite interactions in sticklebacks can lead to a wide range of changes in fish behaviour, colouration, and reproduction (Barber, 2013, Milinski and Bakker, 1990), yet host-parasite-microbe interactions in the three-spined stickleback are only just beginning to be explored. Ling *et al.* (2020) found that different host genotypes differed in their degree of gut microbiome response to infection by a common internal parasite, *Schistocephalus solidus*, suggesting genotype-by-environmental interactions. Despite this emerging evidence of host-parasite-microbe interactions, studies have to date over-simplified the number of interactions to a single parasite species or single parasite exposure. This is not representative of natural systems where multiple parasite infections will occur simultaneously within an individual.

Other studies on the determinants of the wild stickleback gut microbiome identify fish genetics, sex, immune response and polymorphism in the MHC gene as correlated with microbial diversity and community structure (Bolnick *et al.*, 2014a, Bolnick *et al.*, 2014c, Milligan-Myhre *et al.*, 2016, Small *et al.*, 2017, Steury *et al.*, 2019). Additionally, microbial communities have repeatedly shifted in similar directions, with the parallel evolution of sympatric benthic-limnetic pairs from three lakes in Canada, showing that

benthic ecotypes have similar microbial communities across all three lakes, as do limnetic ecotypes (Rennison *et al.*, 2019a). These findings suggest that the gut microbiome could confer a fitness advantage within an ecotype's habitat and may play a key role in adaptation. A crucial tool for elucidating the complicated nature of local adaptation is the reciprocal common garden experiment, where individuals from different populations are placed into a common environment and fitness traits are measured (Kawecki and Ebert, 2004). The strength of these studies comes from the ability to measure genotype-by-environment effects, in the form of both the 'local vs. foreign' hypothesis as well as 'home vs. away' (de Villemereuil *et al.*, 2016, Hoban *et al.*, 2016, Kawecki and Ebert, 2004, Savolainen *et al.*, 2013). These hypothesis state that the resident genotype will have increased fitness in comparison to the non-resident (Kawecki and Ebert, 2004). Local adaptation of the three-spined stickleback has been shown through several common garden experiments (DeFaveri and Merila, 2014, Eizaguirre *et al.*, 2012a, Hendry *et al.*, 2011, Svanback and Schluter, 2012), however, none have yet explored how the gut microbiome may differ within locally-adapted populations. We hypothesis two possible outcomes concerning the coevolution of the three-spined stickleback and their microbiome, as the ecotypes evolve in parallel, each ecotype could show convergence of the microbiome suggesting parallel evolution, or if fish show population-specific microbiomes this would suggest local adaptation.

Due to difficulties in generalising findings across systems, we expanded our research questions to a second model organism. Indeed, if host-microbiome interactions are involved in facilitating the evolution of local adaptation in the host, patterns of population-specific microbiomes should be replicable across systems. Here, we took advantage of access to the third largest aggregation of loggerhead sea turtles globally to test for pattern of host and microbe local adaptation.

1.7.2 The loggerhead sea turtle

Figure 1.5: An image of a female loggerhead sea turtle, Caretta. Image taken by Adrienne Kerley.

Loggerhead sea turtles, *Caretta* (Figure 1.5), are good, but non-classic, model organisms with which to explore host-microbe interactions and their association to local adaptation. They are classified as "Vulnerable" by IUCN (Casale and Tucker, 2017) which is driven by anthropogenic influences such as poaching (Senko *et al*., 2014, Tomillo *et al*., 2008), human development (Taylor and Cozens, 2010), fisheries by-catch (Senko *et al*., 2014), and pollution (Schuyler *et al*., 2016). We focused on loggerhead sea turtles nesting at the Cabo Verde archipelago, which is the third largest nesting aggregation of this species in the world (Marco *et al.*, 2012). Like all sea turtles, the loggerhead shows high level of philopatry, whereby individuals return to their place of birth to breed and lay their eggs (Figure 1.6, Bowen *et al.*, 2004). Philopatric behaviours reduce gene flow among global nesting aggregations (Baltazar-Soares *et al.*, 2020, Shamblin *et al.*, 2014), and among nesting groups within nesting aggregations (Baltazar-Soares *et al.*, 2020, Stiebens *et al.*, 2013b). It is speculated that turtles reach sexual maturity around 45 years of age, after which they will return to their nesting grounds to reproduce (Laloë *et al*., 2014). When not in their nesting habitat, turtles from Cabo Verde

migrate to and use a large feeding ground off the western-coast of Africa with a dichotomous feeding strategy whereby some turtles forage oceanically, and others in coastal waters (Hawkes *et al.*, 2006, Pikesley *et al.*, 2015). Interestingly, to date no segregation of the feeding strategy has been reported based on the natal island, suggesting turtles use the same feeding niches independently of their nesting area. The loggerhead sea turtle is an opportunistic omnivore, feeding on a range of benthic and pelagic prey such as molluscs, siphonophores, jellyfish and nudibranchs (Frick *et al.*, 2009, Tomas *et al.*, 2001).
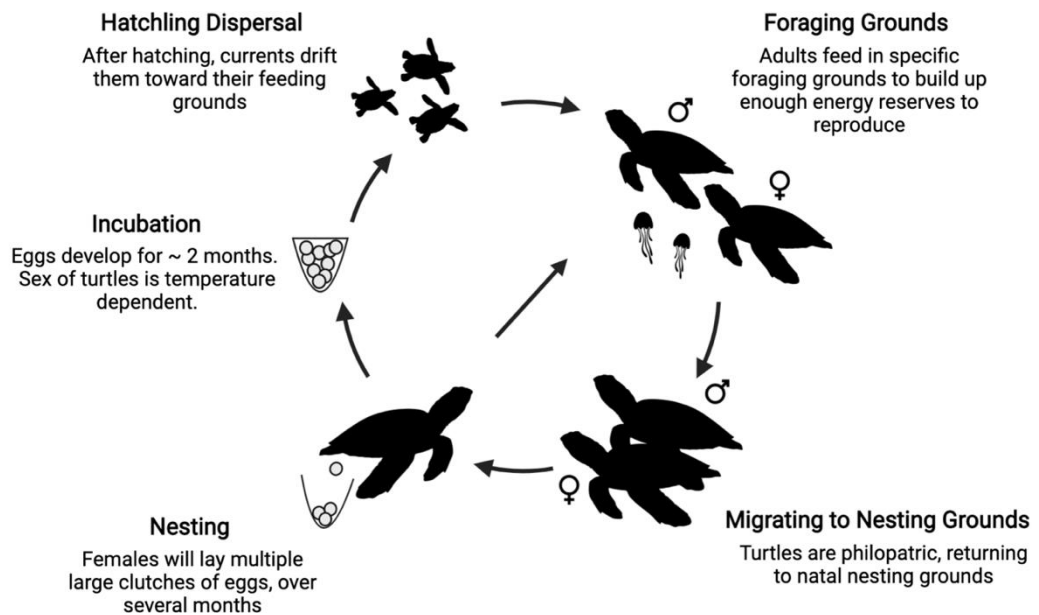
Figure 1.6: A schematic of the lifecycle of a sea turtle. Created in biorender.com

In Cabo Verde, each island supports a nesting group described by unique genetic diversity at neutral mitochondrial and nuclear markers, suggesting population structure in the aggregation (Baltazar-Soares *et al.*, 2020, Stiebens *et al.*, 2013b). It is known there are turtles exploiting two feeding strategies, oceanic and neritic, across all islands, with an additional third neritic feeding strategy found in turtles from Boa Vista (Cameron *et al.*, 2019). Linked to the feeding ecology of loggerhead sea turtles is the infection of the leech parasite, *Ozobranchus margoi*, whereby turtles that have an oceanic feeding

strategy show increased infection prevalence than their neritic counterpart (Lockley *et al.*, 2020). Interestingly, turtles show nesting-group specific (= island-specific) variation in MHC class I genes supporting the idea that high philopatric accuracy has resulted in the evolution of local adaptation (Stiebens *et al.*, 2013a). Because feeding ecology, parasite infection and variation in immune genes have all been linked with changes in the gut microbiome in other organisms, we speculate that locally adapted turtles will harbour island-specific microbial variation across nesting group of the Cabo Verde archipelago (Dheilly, 2014, Muegge *et al.*, 2011, Steury *et al.*, 2019, Sullam *et al.*, 2012).

Most studies on the microbiome of sea turtles have focused on individuals held in rehabilitation centres, which prevent testing for local adaptation in natural environments (Abdelrhman *et al.*, 2016, Biagi *et al.*, 2019, Bloodgood *et al.*, 2020). The impact of rehabilitation has been tested on a small number of green turtles, *Chelonia mydas* (N = 12), which showed wild individuals had higher cloacal microbial diversity than those within the rehabilitation centre, and microbial communities differed between the two groups, potentially suggesting dysbiosis cause by illness or as a result of captivity (Ahasan *et al.*, 2017). Care must be taken however in interpreting results from individuals that are sick or in recovery, as their microbiomes could potentially differ from their healthy counterparts. The first study characterising the loggerhead gut microbiome of turtles, held at a rehabilitation centre, was carried out in 2016, demonstrating how the field is still in its infancy (Abdelrhman *et al.*, 2016). This study highlighted the dominance of Firmicutes, Proteobacteria, Bacteroidetes in both the faecal and intestinal samples of eight individuals, with no difference in microbial community composition between sample types (Abdelrhman *et al.*, 2016). A study on Mediterranean loggerhead sea turtles showed different sizes harboured different microbial communities (Biagi *et al.*, 2019). Two studies have identified the presence of antibiotic resistant bacteria, such as *Citrobacter freundii*, *Proteus vulgaris*, *Providencia rettgeri* and *Pseudomonas*

aeruginosa, within the mouth and cloaca of loggerhead turtles using culture-dependent techniques which likely originated from polluted effluents (Foti *et al.*, 2009, Pace *et al.*, 2019).

A specific study on the gut microbiome of wild loggerhead sea turtles from the USA and Australia showed higher microbial diversity in Australian loggerhead sea turtles, and variation between their microbial communities (Scheelings *et al.*, 2020). However, the small sample size ($N_{USA}= 6$, $N_{Aus}= 18$) may not capture a population's complete microbiome diversity and community structure as hosts exhibit a large amount of inter-individual microbial variation (Bolnick *et al.*, 2014a, Star *et al.*, 2013). Furthermore, given the lack of gene flow and non-overlapping ecology, the comparison lacks relevance. Large scale studies covering a large geographical range and containing a larger number of individuals than previous research are required to sensibly explore what factors might be influencing the loggerhead microbiome and if it is associated with host evolution.

1.8 Thesis outline

This thesis combines chapters on host-microbe and host-parasite interactions placed into an evolutionary context. Microbiology and parasitology were used to understand the selection pressures acting on wild vertebrate populations to evolve local adaptation. I also focused on a host diet as a functional ecological link between parasites and microbes.

Chapter 1 provides a general introduction to the topics covered throughout this thesis and highlights the underlying rational for focusing on two vertebrate species with different life histories.

In *Chapter 2*, I questioned whether locally adapted three-spined stickleback (*Gasterosteus aculeatus*) also showed locally adapted gut microbiomes. Specifically,

with my collaborators, we hypothesised that if stickleback and their microbiome were locally adapted, we would observe population-specific microbiomes. I first compared the gut microbiomes of 11 stickleback populations across Europe and North America to identify whether each population harboured distinct microbiomes. I then explored whether the observed differences were correlated to variation in host genetics, host local environment or a combination of the two.

In *Chapter 3*, I explored the relationship between parasite infection and feeding ecology, both of which are known to influence host-microbe interactions. We performed a field-based reciprocal common garden experiment using three-spined stickleback. Fish were placed into mesocosms in a natural environment, either in a lake or in a river, leaving them to feed and being infected by parasites naturally. After 10 months, fish were screened for macroparasites and muscle samples collected for stable isotope analysis to be used as a long-term proxy of feeding ecology.

In *Chapter 4*, I tested for genotype-by-environment interactions influencing the three-spined stickleback microbiomes. We conducted a field-based reciprocal common garden experiment which allowed us to test for both genetic and environmental influences acting upon the stickleback microbiome. I assessed the microbial diversity and composition of the stickleback gut, their parasite loads as well as their feeding ecology through stable isotope analysis. This allowed for complex host-parasite-microbe and host-microbe diet interactions to be elucidated, which would not have been possible under laboratory conditions.

In *Chapter 5*, I tested whether the patterns I observed in Chapters 2 and 4 could be generalised to other systems, by applying these questions to the loggerhead sea turtle,

(*Caretta caretta*), a highly philopatric species with strong patterns of local adaptation. I focused on turtles from four different nesting islands in the Cabo Verde archipelago. I collected cloacal swab samples to test whether locally adapted nesting groups of turtles also showed population-specific microbiomes. Additionally, we tested for host-parasite-microbe interactions, focusing on the cloacal leech parasite, *Ozobranchus margoi*.

Lastly, in the *General conclusions* chapter, I summarised the overall advances my research contributed to, also providing reflection on how to move forward on the new research avenues emerging from this work.

Chapter 2: Patterns of local adaptation in the microbiome of three-spined sticklebacks across geographical scales

Adrienne Kerley[1], Naga Apoorva Kasala[2], Martin Kalbe[2], Lee Henry[1], China Hanson[3], Manfred Milinski[2] and Christophe Eizaguirre[1]

[1]School of Biological and Chemical Sciences, Queen Mary University of London, London, United Kingdom.

[2]Max Planck Institute for Evolutionary Biology, Department of Evolutionary Ecology, August-Thienemann-Str. 2, 24306 Ploen, Germany.

[3] Microbiology@UCL, University College London, London WC1E 6BT, UK

2.1 Abstract

Symbiotic microbes affect their hosts' physiology, life history traits and even their evolutionary trajectory. Since the last glaciation, the three-spined stickleback (*Gasterosteus aculeatus*) has repeatedly colonised freshwater systems from marine environments, resulting in the evolution of genetically differentiated populations of fish – referred to as ecotypes - in lake and river habitats. This natural experiment allows for the investigation of whether fish belonging to each ecotype have either converged on similar communities of gut microbes (parallel evolution) or are population-specific (local adaptation). To address this question, we compared the gut microbiomes of *G. aculeatus* fish from 11 populations in Europe and North America. We defined operational taxonomic units (OTUs) based on amplicon sequencing of the V3 hypervariable region of the bacterial 16S ribosomal RNA gene. Across all populations, we identified 217 distinct OTUs from 13 bacterial phyla. A total of 93 OTUs were shared across all ecotypes, whilst 22, 12 and 10 OTUs were unique to river, lake, and marine populations respectively. We identify 18 core OTUs (present in ≥50% individuals) across all ecotypes. However, bacterial diversity and community structure, at both the OTU and phylum level, showed that fish microbiomes are best described by the population of origin, rather than ecotype, with few bacterial OTUs shared among populations. This result suggests population-specific microbiomes potentially the result of local adaptation or neutral processes, with a strong environmental determinant of host microbiomes but also some genetic contribution. Together, our results demonstrate that microbial communities are population-specific and that host-microbe interactions are context-dependent in natural populations of stickleback.

2.2 Introduction

Microbiomes play a crucial role in both the ecology and evolution of their hosts (McFall-Ngai *et al.*, 2013). Symbiotic bacteria can provide hosts with benefits such as improved nutrient uptake, regulation of host immune function and pathogen defence (Kelly and Salinas, 2017, Lawley and Walker, 2013, Lee and Hase, 2014). Given their functional importance, microbiomes have also been implicated in the rapid ecological adaptation of their hosts (Alberdi *et al.*, 2016). Advances in our understanding of vertebrate microbiomes show that both environmental and genetic factors determine the diversity and structure of a host's gut microbiome (Grube *et al.*, 2012, Ley *et al.*, 2008a). However, our understanding of the relative importance of host genetics versus environmental factors in shaping the composition of vertebrate microbiomes is limited.

Direct environmental factors such as temperature and humidity, as well as salinity, residence time or pH in the aquatic realm determine the diversity of bacteria a host is exposed to, and in turn, colonised by (Apprill, 2017, Krause *et al.*, 2012, Lindström and Bergström, 2004, Lozupone and Knight, 2007). Ecological assemblages also affect the composition of prey and parasites a host is exposed to, both of which indirectly impact gut microbiota (David *et al.*, 2014, Ling *et al.*, 2020, Rausch *et al.*, 2013). Different food sources not only carry their own specific microbiomes, but also differ in the nutrients they contain, which will be assimilated by different bacterial species within the host (Desai *et al.*, 2012, Faith *et al.*, 2011, Hildebrandt *et al.*, 2009, Parks *et al.*, 2013, Turnbaugh *et al.*, 2009, Wu *et al.*, 2011). For instance, diet strongly contributes to the evolution of the gut microbiome of ant-eating mammals, with convergence of microbial diversity and composition across species and global distribution (Delsuc *et al.*, 2014). These results show that diet adaptation is a major factor of gut microbiome composition over evolutionary timescales.

In addition to environmental determinism, a host's genotype can also influence the composition of its microbiome (Smith *et al.*, 2015, Steury *et al.*, 2019, Sullam *et al.*, 2012, Wang *et al.*, 2016a). Genetic control can stem from host immunity and immunity-related genes (Bates *et al.*, 2007). Variation at important immunity loci, such as within the major histocompatibility complex (MHC) gene region, are known to be associated with the composition of the human microbiome (Bonder *et al.*, 2016). Immunity-based selection by the host favours genotype-specific microbial communities that are stable with a high level of functional redundancy (Ley *et al.*, 2006). Closely related individuals are shown to harbour more similar microbiomes than unrelated subjects exposed to similar environments (Chen *et al.*, 2018, Zoetendal *et al.*, 2001). In some organisms, host genotype can override the impact environment has on the gut microbiome - a particularly powerful example comes from two *Hydra* species, that retained distinct microbial communities after two decades under standardized laboratory conditions. Even more striking is these communities remained similar to their wild counterparts (Fraune and Bosch, 2007). In species where sex is genetically determined, a host's genetics and environment can interact to influence their microbiome, for instance in fish, sex-specific diets result in different gut microbiomes between the sexes (Bolnick *et al.*, 2014c). In humans and mice, the interaction between dietary fibre intake and sex results in changes in gut microbiome diversity and composition (Dominianni *et al.*, 2015, Zhang *et al.*, 2020).

It remains to be tested in nature whether populations adapted to similar habitats have converged on similar microbiomes. If specific conditions exist across similar, but geographically distinct locations, parallel evolution of certain traits can emerge (Colosimo *et al.*, 2005, Elmer *et al.*, 2010, Marchinko and Schluter, 2007). Yet, a common outcome of evolution is the local adaptation of host populations (Eizaguirre *et al.*, 2012a, Savolainen *et al.*, 2013, Sobel *et al.*, 2010). In this context, if a host

population's microbiome were locally adapted, we would expect to see population-specific microbiomes, with few OTUs shared between populations. On the other hand, if parallel evolution of the microbiome has occurred, we would expect to see similar microbiomes between host lineages that have adapted to the same habitats. To test such a hypothesis requires an organism that has repeatedly colonised and adapted to different habitats.

The three-spined stickleback (*Gasterosteus aculeatus*) provides a unique opportunity to disentangle the factors contributing to the composition of vertebrate microbiomes. Freshwater fish, referred to as lake and river ecotypes, have repeatedly evolved from marine and anadromous populations across the northern hemisphere (Bell and Foster, 1994, Berner *et al.*, 2009, Eizaguirre *et al.*, 2011, Feulner *et al.*, 2015, Marques *et al.*, 2016, Ravinet *et al.*, 2013, Reusch *et al.*, 2001). Within a lake-river pair, ecotypes show a large amount of morphological (Cano *et al.*, 2006), physiological (Taylor and McPhail, 2000) and genetic differentiation (Feulner *et al.*, 2015, Marques *et al.*, 2016), which have derived from parallel and non-parallel evolution (Bolnick *et al.*, 2018, Feulner *et al.*, 2015, Hanson *et al.*, 2017). The gut microbiome of North American three-spined stickleback populations varies with diet, sex, ecotype, habitat geomorphology, and polymorphism in the adaptive immune genes of the major histocompatibility complex (Bolnick *et al.*, 2014a, Bolnick *et al.*, 2014b, Bolnick *et al.*, 2014c, Smith *et al.*, 2015). A recent study suggested that, in the three-spined stickleback, microbiome diversity and composition were more strongly correlated to a population's genetic divergence than to host environment or geography (Steury *et al.*, 2019). However, whether fish lineages that have repeatedly colonised the same habitats have converged on a similar gut microbiome has yet to be tested.

Here, we use the naturally replicated evolution of three-spined stickleback ecotypes, which have undergone parallel divergence across multiple locations

independently, from a common ancestor, to test how the microbiome of fish from different wild populations and ecotypes vary to identify if this ecotype divergence is associated with changes in the microbiota. Specifically, we test the hypothesis that; i) similar microbial communities will be shared among individuals within an ecotype across different sampling sites, suggesting parallel evolution, or alternatively, or ii) individuals from the same population will harbour population-specific microbiomes, which may demonstrate local adaptation. Although such patterns may also be driven by neutral processes or microbiota demonstrating ecological flexibility.

2.3 Materials and Methods

2.3.1 Sample Collection

A total of 178 three-spined sticklebacks were collected from 11 locations across Europe and North America (Table 2.1). Distances between sampling sites can be found in Table 2.2. Three lake-river ecotype pairs from Europe ($N_{population} = 6$) and one North American ecotype pair ($N_{population} = 2$), one lake ecotype from Greenland ($N_{population} = 1$), and two marine ecotypes from Germany and Canada ($N_{population} = 2$). Asymmetry in sampling was due to access to certain locations and availability of three-spined sticklebacks at sampling sites. Fish were caught using a hand net. European sticklebacks were euthanised and dissected in the laboratory within three days of capture. North American sticklebacks were euthanised, frozen on-site and dissected later. Fish standard length, weight, sex, spleen, and liver weight were recorded. Whole intestines were dissected using aseptic techniques, weighed, and stored in RNAlater® at -80°C. Using the entire gut allowed us to characterize residential microbiota, found in the intestinal epithelium and transient bacteria, located in the gut lumen.

Table 2.1: Sample site information with continent, country, ecotype classification, coordinates, number of individual sticklebacks sequenced and retained per population after rarefaction to 3000 reads per fish gut sample, N = 178 fish intestine samples sequenced, and N = 157 fish intestine samples retained post-rarefaction.

Continent	Country	Population	Ecotype	No˚ of fish sequenced	No˚ of fish post-rarefaction	Latitude	Longitude
Europe	Germany	Großer Plöner See	Lake	16	15	54˚09'21.61"N	10˚25'48.52"E
		Malenter Au	River	16	13	54˚12'15.08"N	10˚33'41.90"E
		Westensee	Lake	16	15	54˚17'01.92"N	9˚56'55.71"E
		Eider	River	16	14	54˚18'12.2"N	9˚57'16.4"E
		Fehmarn	Marine	16	15	54˚28'55.2"N	11˚00'36.5"E
	Norway	Skogseidvatnet River	River	28	25	60˚15'15.05"N	5˚55'29.28"E
		Skogseidvatnet Lake	Lake	18	15	60˚14'41.57"N	5˚54'55.39"E
North America	Canada	Brannen Lake	Lake	12	8	49˚12'43.08"N	124˚03'44.83"W
		Millstone River	River	12	11	49˚10'34.35"N	123˚57'45.19"W
		Millstone Estuary	Marine	12	10	49˚10'16.60"N	123˚6'12.16"W
	Greenland	Badesø Lake	Lake	16	16	64˚07'49.78"N	51˚22'21.11"W

Table 2.2:Distances in kilometres between sampling sites.

	Großer Plöner See	Malenter Au	Westensee	Eider	Fehmarn	Skogseidvatnet River	Skogseidvatnet Lake	Brannen Lake	Millstone River	Millstone Estuary	Badesø Lake
Großer Plöner See	-	10.1	34.4	35.0	52.2	730.0	729.2	7759.0	7759.0	7732.0	3533.0
Malenter Au	10.1	-	40.8	41.0	42.4	727.9	727.2	7759.0	7759.0	7731.0	3536.0
Westensee	34.4	40.8	-	2.2	72.2	706.3	705.5	7731.0	7731.0	7704.0	3501.0
Eider	35.0	41.0	2.2	-	71.2	704.3	703.5	7730.0	7730.0	7702.0	3500.0
Fehmarn	52.2	42.4	72.2	71.2	-	710.1	709.4	7749.0	7746.0	7719.0	3536.0
Skogseidvatnet River	730.0	727.9	706.3	704.3	710.1	-	1.2	7040.0	7040.0	7015.0	2900.0
Skogseidvatnet Lake	729.2	727.2	705.5	703.5	709.4	1.2	-	7041.0	7041.0	7015.0	2900.0
Brannen Lake	7759.0	7759.0	7731.0	7730.0	7746.0	7040.0	7041.0	-	8.3	69.8	4448.0
Millstone River	7759.0	7759.0	7731.0	7730.0	7746.0	7040.0	7041.0	8.3	-	62.5	4446.0
Millstone Estuary	7732.0	7731.0	7704.0	7702.0	7719.0	7015.0	7015.0	69.8	62.5	-	4406.0
Badesø Lake	3533.0	3536.0	3501.0	3500.0	3536.0	2900.0	2900.0	4448.0	4446.0	4406.0	-

2.3.2 Extraction, amplification, and sequencing

DNA was extracted from whole intestines using QIAamp DNA Stool Mini kit (Qiagen) following manufacturer's protocol. Amplification of the V3 hypervariable region (152 – 197bp) of the 16S ribosomal RNA gene was carried out using the forward primer 341F (5′-TCCTACGGGNGGCWGCAG-3') and the reverse primer 785R (5'-TGACTACHVGGGTATCTAAKCC-3') (Klindworth *et al.*, 2013). The V3 region was chosen as it is well documented in most reference databases, covers a broad spectrum of microbial diversity and is a single region saving on cost and sequencing time (Garcia-Lopez et al., 2020). 16S rRNA was amplified using 2 µl 5x MyTaq Reaction Buffer, 1.5 units MyTaq DNA polymerase (Bioline), 2 µl BioStabII PCR enhancer (Sigma), 5 ng template DNA, and 1.5 µl of forward and reverse primer (10 pmol/µl), per 20 µl-volume reaction. Each individual was marked with a unique barcode at the 5' end prior to pooling. PCRs were conducted using the following protocol: 2 minutes at 96°C, 15 seconds at 96°C (x30), 20 seconds at 50°C (x30), and 60 seconds at 72°C (x30). Amplicons were sent to LGC Genomics (Berlin, Germany) for Illumina MiSeq 300bp paired-end sequencing. Samples were retained if barcodes matched entirely. Reads longer than 100bp were retained and primers were allowed 2 mismatches.

2.3.3 Data analyses

Demultiplexing, adaptor, primer and quality trimming were completed using Illumina's CASAVA, TruSeq™, and FLASh 1.2.4 (Supplementary Information 2.1). Chimeras were detected and removed, using USEARCH 6.1 (Edgar, 2010). Clustering of sequences into Operational taxonomic units (OTUs) was carried out in QIIME 1.9.1 (Caporaso *et al.*, 2010). OTUs were identified using open-reference OTU picking, using the USEARCH algorithm and QIIME defaults at 97% as well as 99% similarity threshold. A more conservative threshold was used to allow for the possibility of more subtle structure (Chen *et al.*, 2013) linked to the recent divergence between some fish ecotypes

(Berner *et al.*, 2008, Reusch *et al.*, 2001). OTUs were taxonomically classified using Greengenes 13_8 release (DeSantis *et al.*, 2006). Sequences were rarefied to 3000 reads to standardise the sampling effort (Smith *et al.*, 2015). Statistical analyses were conducted in R version 3.6.3, packages used include: *phyloseq* (McMurdie and Holmes, 2013), *vegan, lmerTest, lme4* (R Core Team, 2020). Mitochondria and chloroplast reads were removed as well as unassigned OTUs at the kingdom level. For community composition analyses, further filtering was undertaken to remove low abundance taxa, defined as OTUs with fewer than 10 reads across at least 2 samples. For microbial diversity metrics, these taxa were retained. Normality of model residuals were tested and transformed if required to meet test assumptions. All linear models (LM) were backward selected using Akaike's Information Criterion (AIC) values to retain the optimal reduced model.

Good's coverage was calculated in QIIME to estimate sample completeness. Understanding whether similar bacteria are common across all populations, across ecotypes or whether they are population specific will define whether host-symbiont interactions evolve in parallel or are locally adapted. To this end, we identified OTUs that were ubiquitous, partially shared or unique to a given ecotype. Core bacteria were identified as OTUs present in ≥50% of the overall fish gut samples but present in at least one fish per population (Sullam *et al.*, 2015) across all samples and within ecotypes.

Fish microbial diversity (alpha diversity) was calculated at the OTU and phylum level, using Shannon's diversity index. Linear models were used to test for correlation between Shannon's diversity indices and continent, sex, ecotype, standard length, population nested within continent, and their interactions. Due to covariance between fish standard length, continent and population of origin, residuals of a linear model between the variables were used where necessary.

Using PERMANOVAs based on Bray-Curtis, weighted UniFrac and unweighted UniFrac distance matrices (beta diversity), we tested whether OTU and phylum

composition were associated with continent, ecotype, standard length of fish (grouped into small, medium and large size categories per population), populations nested within continent, sex, and their interactions. Post-hoc tests were then performed on significant variables using Bonferroni corrections. In order to test for potential parallel evolution of microbial communities, PERMANOVAs on subsets of lake ecotype populations and river ecotype populations were performed at OTU and phylum levels with the same variables as mentioned above. To investigate microbial community structure of ecotype pairs, we used PERMANOVAs, at OTU and phylum levels, based on Bray-Curtis, weighted UniFrac and unweighted UniFrac distance matrices, to compare microbial community composition between the four lake-river ecotype pairs, German ecotype pair 1: Westensee - Eider, German ecotype pair 2: Großer Plöner See – Malenter Au, a Norwegian ecotype pair: Skogseidvatnet lake - Skogseidvatnet river and a Canadian ecotype pair: Brannen lake - Millstone river. Furthermore, to decipher differences in community structure of within lake-river ecotype pairs, PERMANOVAs were carried out on the within the four lake-river pairs. Similarity percentages (SIMPERs) were used to identify the OTUs/phyla explaining most variation between each lake-river ecotype pair. Mantel tests were used to test for correlations between pairwise geographic distances and metrics of stickleback gut microbiome differentiations obtained from the PERMANOVAs to test for isolation by distance.

We used linear models to test for associations between traits thought to be important for fish fitness and fish populations. We used standard length (SL), body condition (CF $= \left(\frac{Body\ Weight}{Standard\ Length^{2.89}} \right) \times 100$) (Frischknecht, 1993), splenosomatic (SSI $= \frac{Spleen\ Weight}{Body\ Weight} \times 100$) and hepatosomatic indices (HSI $= \frac{Liver\ Weight}{Body\ Weight} \times 100$) as fish fitness traits. Lastly, we tested whether fitness traits were correlated with an interaction between the OTUs highlighted as varying in abundance with ecotype pairs in the SIMPER analysis and ecotype pair populations using linear models.

2.4 Results

A total of 1,082,886 reads were obtained from 178 individual gut microbiome samples, with an average of 6,084 (SD± 2,225) reads per sample. Rarefaction to 3000 reads based on 97% clusters resulted in 16 samples being removed, additionally, fish who were not assigned a sex (N = 5) were also removed, leaving a total of 157 individual samples. After rarefaction and removing unassigned kingdoms, mitochondria and chloroplasts, we retrieved 4,622 bacteria OTUs that were converted into Shannon's diversity index. Singletons and low abundance OTUs were removed (fewer than 10 reads across at least 2 samples), resulting in a total of 217 OTUs across all samples, which were used for all community-based analyses. Good's coverage values for all fish were ≥ 96%, indicating that the majority of microbial species were accounted for in each sample (Supplementary Table 2.1). A total of 13 bacterial phyla were detected, with 4 phyla accounting for ~93% of the reads (Supplementary Table 2.2). Firmicutes, Proteobacteria and Actinobacteria were common across all fish samples, whereas Spirochaetes were only observed in European samples, yet with a large amount of interindividual variation (Figure 2.1).

OTU clustering at 99% sequence similarity resulted in no significant patterns and hence we focus on reporting results from OTUs and phyla identified at the 97% sequence similarity level (Supplementary Table 2.3 & 2.4).
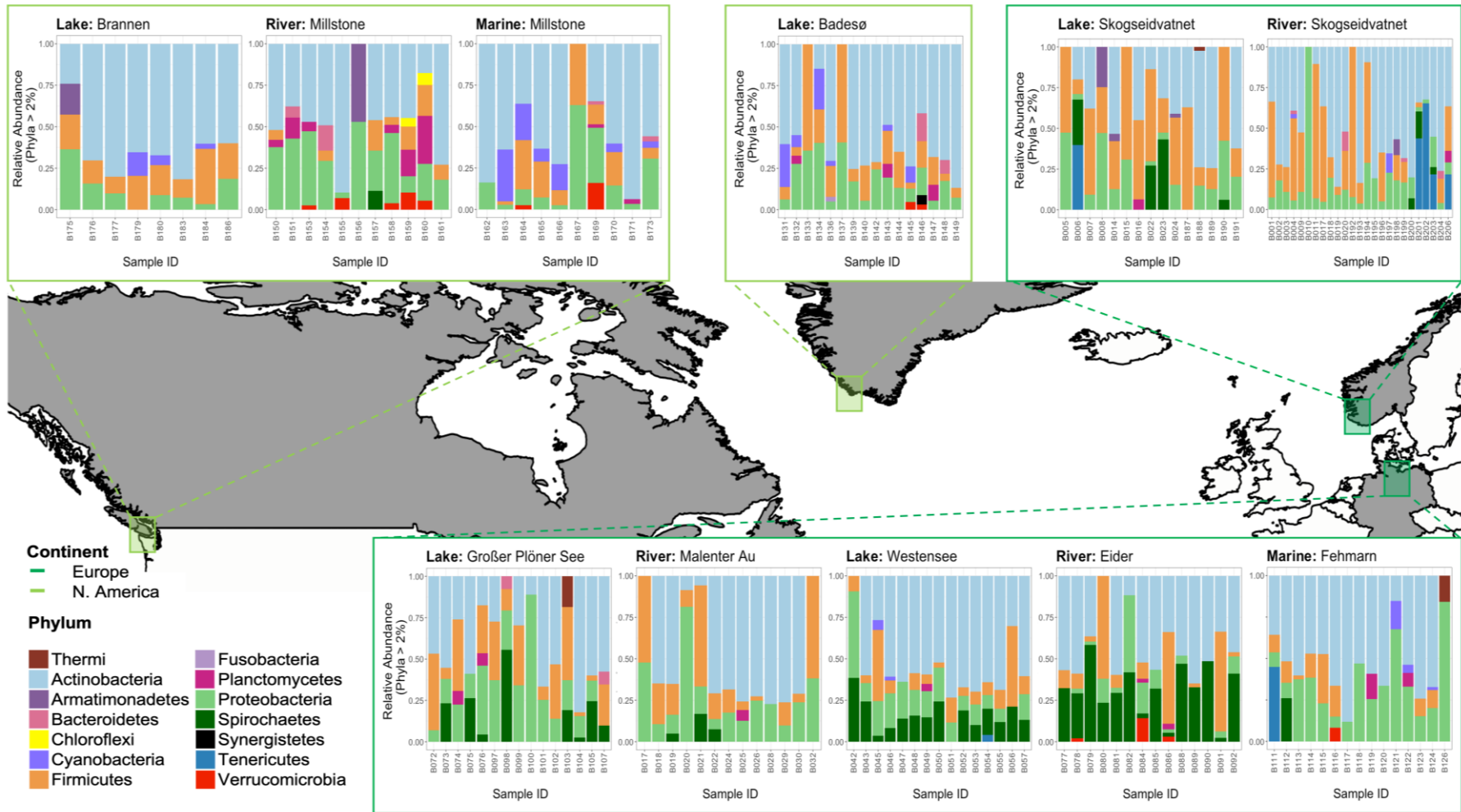
Figure 2.1: Relative abundance of bacterial composition at phylum level per individual for each of the 11 populations of three-spined stickleback sampled. Phyla with abundance <2% per individual are not shown

2.4.1 Operational taxonomic units

Out of the 217 OTUs observed, 93 were shared across marine, lake and river ecotypes. We found 22 OTUs unique to river ecotypes; 12 unique to lake ecotypes, and 10 unique to marine ecotypes. (Figure 2.2A). When looking specifically at OTU abundance, a total of 18 core OTUs were identified in 50% of the samples, in at least one fish per population (Figure 2.2B). We found 17 of the 18 core OTUs were classified as core across all ecotypes, while a *Microbacterium species* (OTU 748636) was a core OTU in lake and marine fish but not within river fish. Within lake fish, an additional OTU, OTU 16121 (*Clostridium* sp.), was classified as core. No core OTUs were unique to river or marine ecotypes.

The majority of phyla found, 9 of 13, were present in all three ecotypes (Figure 2.2C). The phylum Chloroflexi was only observed in river fish, while Fusobacteria was solely in fish collected from lakes. Planctomycetes was found in both lake and river fish but was not present in marine samples. Verrucomicrobia was present in both river and marine fish but absent from lake fish. The number of OTUs and phyla were not randomly distributed across fish ecotype (Chi-squared test, OTU: $\chi^2(432, N = 3) = 136630$, p < 0.001, Phylum: $\chi^2(24, N = 3) = 25405$, p < 0.001).

A

Lake:
177 OTUs

River:
181 OTUs

12   58   22

14   93   8

10

Marine:
125 OTUs

B

Lake:
19 Core OTUs

River:
17 Core OTUs

1   0   0

1   17   0

0

Marine:
18 Core OTUs

C

Lake:
11 Phyla

River:
12 Phyla

1   1   1

0   9   1

0
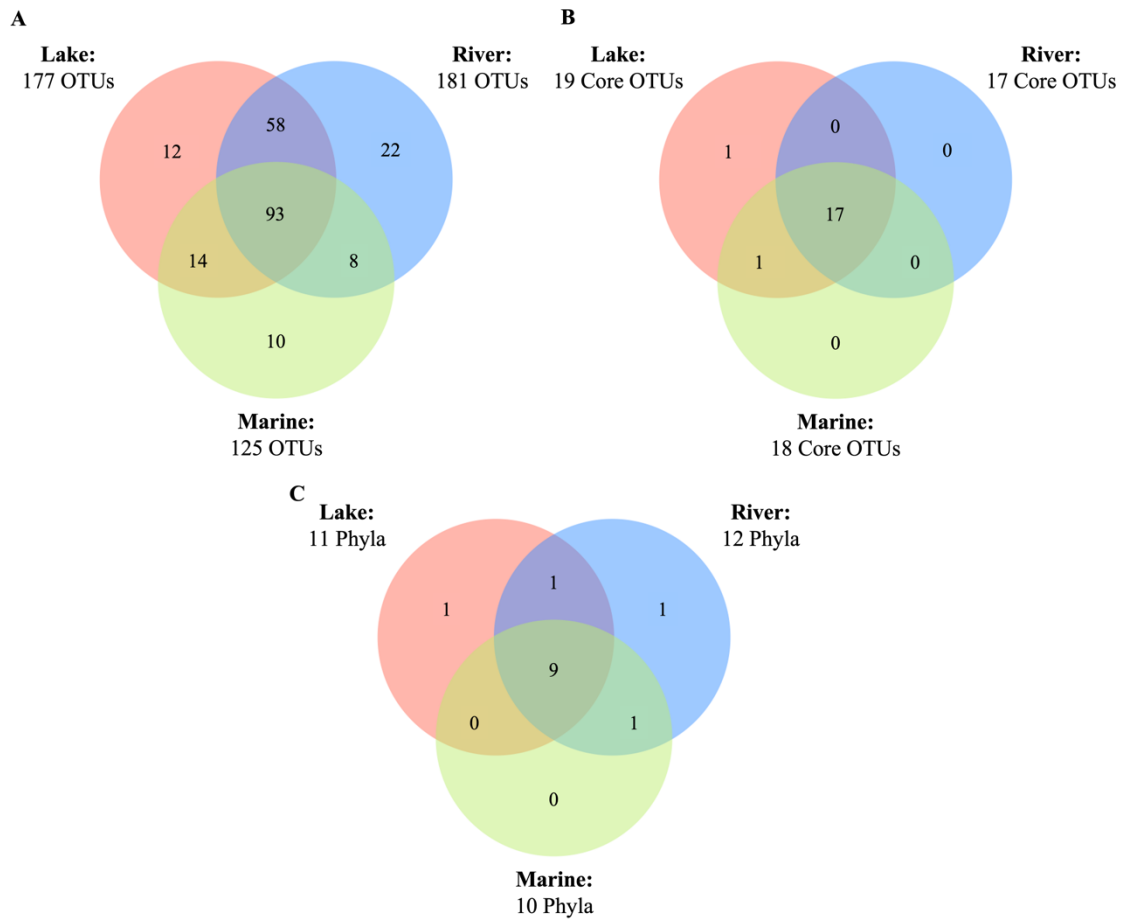
Marine:
10 Phyla

Figure 2.2: A) Shared and distinct OTUs sampled across lake, river and marine ecotype fish, B) Shared and distinct core OTUs (≥50%) detected across lake, river and marine ecotype fish. C) Shared and distinct microbial phyla detected across lake, river and marine ecotype fish. Shared OTUs and phyla were determined by their presence in at least one individual within each ecotype, core OTUs ≥50% of the overall samples but present in at least one fish per population.

2.4.2 Microbial diversity

The Shannon's diversity index of OTUs was negatively correlated with fish length (LM, $F_{1,154} = 4.40$, $p < 0.05$, Table 2). No other variables significantly correlated with OTU diversity. At the phylum level, we also found a negative association between bacterial diversity and fish size (LM, $F_{1,140} = 9.90$, $p < 0.01$, Table 2.3). We found phylum level diversity varied across populations nested within continent (LM, $F_{9,140} = 2.40$, $p < 0.05$, Table 2.3, Figure 2.3). An interaction between fish sex and continent was also detected (LM, $F_{1,140} = 4.09$, $p < 0.05$; Table 2.3, Supplementary Figure 2.1), whereby no differences existed between males and females in European populations, while males

showed reduced Shannon diversity compared to females in North America (Least Square Means: Europe: $p > 0.05$, North America: $p < 0.05$).

Table 2.3: Summary statistics evaluating the effects of continent, population nested within continent, sex, ecotype and standard length on microbial diversity using a linear model at the a) OTU level and b) phylum level. Models were backward selected using the step function. Significant results are highlighted in bold. Df denotes degrees of freedom.

	Df	F	P-value	
(a) OTU level				
Continent	(1,154)	3.27	0.072	.
Standard Length	**(1,154)**	**4.39**	**0.038**	*
(b) Phylum level				
Continent	(1,140)	1.24	0.268	
Sex	(1,140)	0.45	0.504	
Standard Length (Residuals)	**(1,140)**	**9.90**	**0.002**	**
Continent:Population	**(9,140)**	**2.40**	**0.015**	*
Continent:Sex	**(1,140)**	**4.09**	**0.045**	*
Continent:Standard Length (Residuals)	(1,140)	0.49	0.484	
Sex:Standard Length (Residuals)	(1,140)	0.001	0.940	
Continent:Sex:Standard Length (Residuals)	(1,140)	3.62	0.059	.
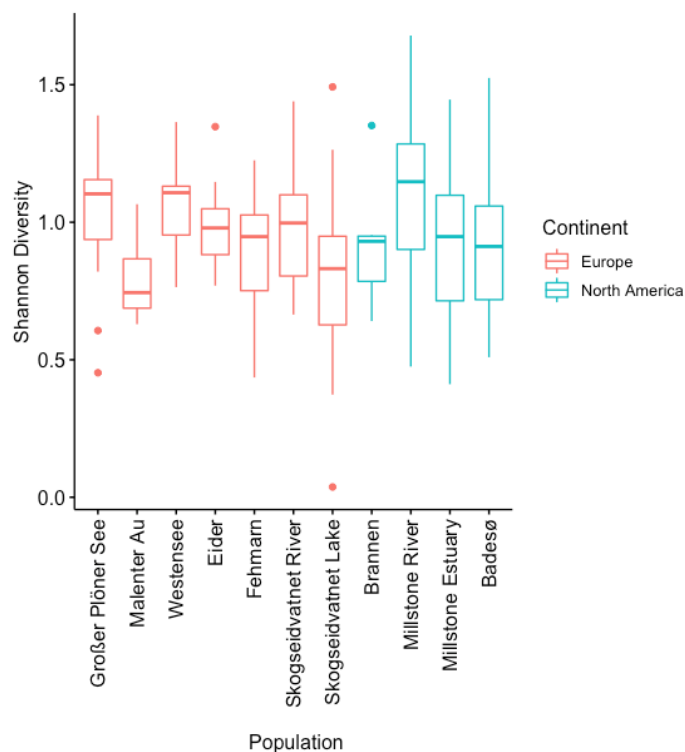
Figure 2.3: Shannon diversity of OTUs at the phylum level varies with an interaction between fish populations nested within each continent (LM: $F_{9,140} = 2.403$, $p < 0.05$).

2.4.3 Stickleback microbiome composition

The composition of intestinal microbial communities based on three different β-diversity metrics revealed consistent results for OTU and phylum levels. Specifically, within each continent, fish harboured population-specific bacterial communities as shown for all three β-diversity metrics, estimated at both OTU and the phylum levels (PERMANOVA, All β-diversity metrics: $p < 0.05$, Table 2.4). We also found continent and ecotype-specific bacterial communities at OTU and phylum level (PERMANOVA, All β-diversity metrics: $p < 0.05$, Table 2.4). Fish from marine and lake ecotypes consistently differed in their microbial composition, across all three β-diversity distances, at both OTU and phylum level (pairwise PERMANOVA, All β-diversity metrics: $p < 0.05$, Supplementary Table 2.5). Marine populations also showed different OTU and phylum communities to those observed in river fish and this was particularly true for unweighted UniFrac (pairwise PERMANOVA, $p < 0.01$, Supplementary Table 2.5). Microbial communities of lake and river fish differ only for unweighted UniFrac at the OTU level (pairwise PERMANOVA, $p < 0.05$, Supplementary Table 2.5). The main difference between distance matrices and phylum or OTU level analysis of microbiomes was observed for fish from different length classes that harboured different microbial communities with weighted UniFrac and Bray-Curtis at the phylum level but at the OTU level the only significant β-diversity metrics was Bray-Curtis (PERMANOVA, Phylum: Weighted UniFrac: $F_{2,156} = 2.68$, $p < 0.01$, and Bray-Curtis: $F_{2,156} = 2.38$, $p < 0.05$, OTU: Bray-Curtis: $F_{2,156} = 1.68$, $p < 0.05$, Table 2.4). We found a pattern of isolation by distance whereby stickleback microbiome community composition increased in differentiation with increasing distance using unweighted UniFrac and Bray-Curtis (Mantel Unweighted UniFrac: $R = 0.418$, $p < 0.05$, Bray-Curtis: $R = 0.448$, $p < 0.05$, but not when using weighted UniFrac: $R = 0.205$, $p = 0.134$, Figure 2.4).

In order to test for parallel evolution of the gut microbiome within an ecotype, we compared microbial communities among populations of lake fish and among populations of river fish respectively. We did not investigate among marine populations as only 2 sites were sampled. At both the OTU and phylum level, we found that fish harboured population-specific bacterial communities within each ecotype, independently of the community composition β-diversity metric used ($p < 0.05$, Table 2.5).

Three-spined sticklebacks have evolved into lake and river ecotype pairs with strong genetic and morphological differences. Our results show that each population pair follows transition-specific dynamics at a number of levels. Firstly, the four lake-river ecotype transitions, two German, one Norwegian and one Canadian pair, differed in microbial communities at the OTU (PERMANOVA, Weighted UniFrac: $F_{3,115} = 4.95$, $p < 0.001$, Unweighted UniFrac: $F_{3,115} = 3.06$, $p < 0.001$ and Bray-Curtis: $F_{3,115} = 4.08$, $p < 0.001$) and phylum level (PERMANOVA, Weighted UniFrac: $F_{3,115} = 6.46$, $p < 0.001$, Unweighted UniFrac: $F_{3,115} = 5.62$, $p < 0.001$ and Bray-Curtis: $F_{3,115} = 6.12$, $p < 0.001$). We then tested whether microbiomes of fish consistently differed within an ecotype pair. Three out of four lake-river ecotype pairs showed fish bacterial communities differ significantly at both OTU and phylum level across the community composition metrics (Supplementary Table 2.6 & Supplementary Figure 2.2). The Großer Plöner See (lake) and Malenter Au (river) were the only population pair where fish carried similar microbial communities (Supplementary Table 2.6 & Supplementary Figure 2.2). Across the four ecotype pairs, we found that 30 OTUs significantly differed in abundance between fish ecotypes. Among those, 4 were common across population-pairs from both continents. All 26 remaining were specific to each lake-river transition. Specifically, 6 OTUs differed between Canadian lake-river fish, 11 OTUs for the Norwegian fish, 5 OTUs for the Westensee - Eider German fish, and 8 OTUs between the Großer Plöner See - Malenter Au fish (Table 2.6).

Table 2.4: Nested PERMANOVA of three β-diversity metrics showing the effect of continent, standard length classes, ecotype and population (which is nested within continent) at OTU and phylum level. Permutations: 1000. Significant results are shown in bold. Df denotes degrees of freedom.

| | Df | *Bray-Curtis* | | | *Weighted UniFrac* | | | *Unweighted UniFrac* | | |
		F	R^2	P-value	F	R^2	P-value	F	R^2	P-value
OTU level										
Continent	(1,156)	4.41	0.02	**<0.001**	5.54	0.03	**<0.001**	4.10	0.02	**<0.001**
Ecotype	(2,156)	2.47	0.03	**<0.01**	2.80	0.03	**<0.01**	3.30	0.04	**<0.001**
Standard Length (Categories)	(2,156)	1.68	0.02	**<0.05**	1.51	0.02	0.113	1.12	0.01	0.275
Continent:Population	(7,156)	2.94	0.12	**<0.001**	3.58	0.14	**<0.001**	2.16	0.09	**<0.001**
Phylum level										
Continent	(1,156)	6.04	0.03	**<0.001**	5.64	0.03	**<0.001**	8.67	0.05	**<0.001**
Ecotype	(2,156)	2.63	0.03	**<0.05**	2.53	0.03	**<0.05**	4.26	0.05	**<0.001**
Standard Length (Categories)	(2,156)	2.38	0.03	**<0.05**	2.68	0.03	**<0.01**	0.75	0.01	0.606
Continent:Population	(7,156)	4.26	0.16	**<0.001**	4.26	0.16	**<0.001**	2.74	0.11	**<0.001**
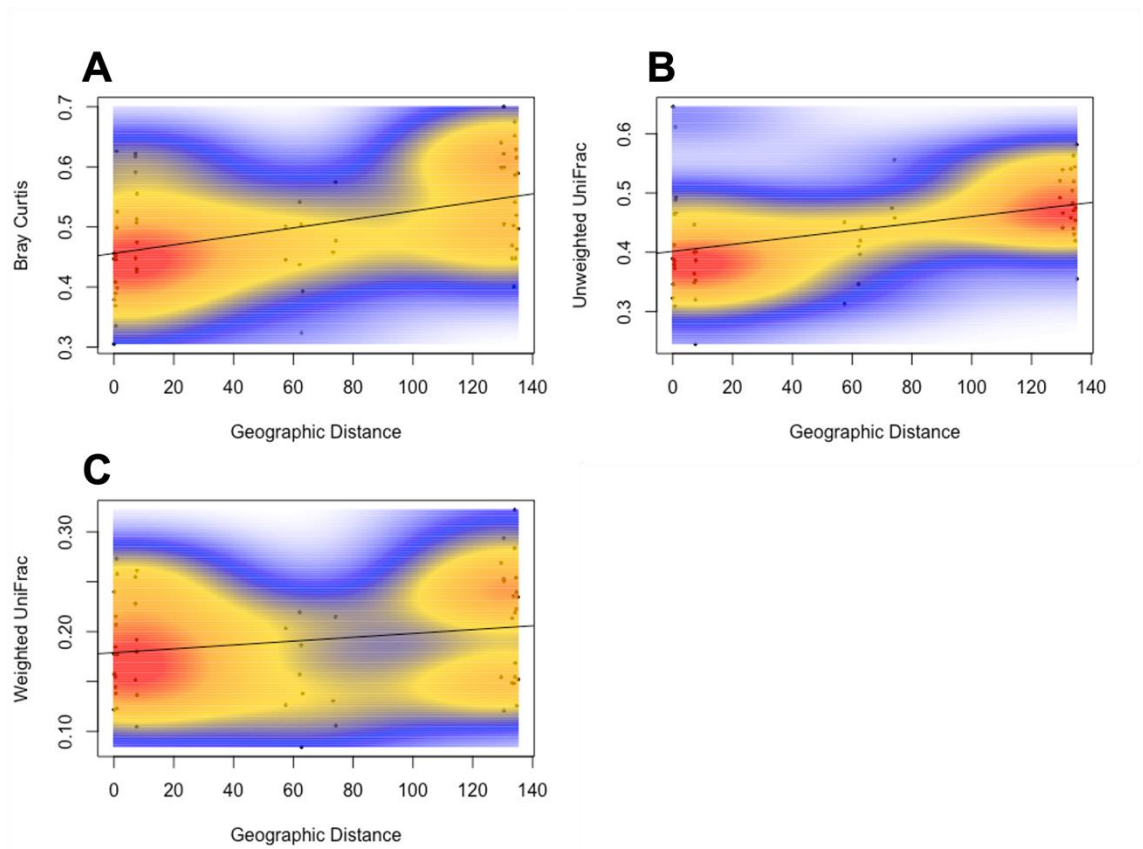
Figure 2.4: Mantel test for isolation by distance between the matrix of geographic distances and A) Bray Curtis (R = 0.448, p < 0.05), B) Unweighted UniFrac (R = 0.418, p < 0.05) and C) Weighted UniFrac (R = 0.205, p = 0.134). Colours represent the relative density of points, with red showing higher densities and blue lower densities, line shows the correlation between the two distance matrices.

Table 2.5: PERMANOVA of three β-diversity metrics among lake populations and river populations showing the effect of population, at OTU and phylum level. Permutations:1000. Significant results are shown in bold Df denotes degrees of freedom.

	Df	F	R^2	P-value	
Lake Populations					
OTU level					
Bray-Curtis	**(4,68)**	**2.81**	**0.15**	**<0.001**	***
Weighted UniFrac	**(4,68)**	**3.00**	**0.16**	**<0.001**	***
Unweighted UniFrac	**(4,68)**	**1.83**	**0.10**	**<0.001**	***
Phylum level					
Bray-Curtis	**(4,68)**	**4.32**	**0.21**	**<0.001**	***
Weighted UniFrac	**(4,68)**	**3.73**	**0.19**	**<0.001**	***
Unweighted UniFrac	**(4,68)**	**2.56**	**0.14**	**0.007**	**
River Populations					
OTU level					
Bray-Curtis	**(3,62)**	**3.27**	**0.14**	**<0.001**	***
Weighted UniFrac	**(3,62)**	**5.07**	**0.20**	**<0.001**	***
Unweighted UniFrac	**(3,62)**	**3.06**	**0.13**	**<0.001**	***
Phylum level					
Bray-Curtis	**(3,62)**	**4.85**	**0.20**	**<0.001**	***
Weighted UniFrac	**(3,62)**	**5.65**	**0.22**	**<0.001**	***
Unweighted UniFrac	**(3,62)**	**4.89**	**0.20**	**<0.001**	***

Table 2.6: OTUs found in significantly in abundance among lake-river ecotype pairs derived from SIMPER analysis.

		Lake-River Ecotype pair			
OTU	Taxonomic information	Canadian	Norwegian	German Ecotype Pair 1	German Ecotype Pair 2
4342297	*Clostridium* sp.	✔	-	✔	-
288283	*Caulobacteraceae family*	✔	-	-	-
16121	*Clostridium* sp.	✔	-	-	✔
73202	*Clostridium* sp.	✔	-	-	-
53	*Clostridium* sp.	✔	-	-	✔
54732	*Propionibacterium acnes*	✔	-	-	-
24	Brevinemataceae family	-	✔	-	-
1088265	*Propionibacterium acnes*	-	✔	-	-
848816	Bacillaceae family	-	✔	-	-
4315319	*Deefgea* sp.	-	✔	-	-
697578	Bacillaceae family	-	✔	-	-
207	Bacillaceae family	-	✔	-	-
875118	*Propionibacterium acnes*	-	✔	-	-
839235	Aeromonadaceae family	-	✔	-	-
94906	*Bacillus humi*	-	✔	-	-
814133	*Synechococcus* sp.	-	✔	-	-
103	Bacillaceae family	-	✔	-	-
95	*Spironema* sp.	-	-	✔	-
112057	*Renibacterium* sp.	-	-	✔	-
3202924	*Propionibacterium acnes*	-	-	✔	-
162	Endozoicimonaceae family	-	-	✔	✔
101445	Methylophilaceae family	-	-	-	✔
1087597	*Propionibacterium acnes*	-	-	-	✔
81821	Methylophilaceae family	-	-	-	✔
403853	*Propionibacterium acnes*	-	-	-	✔
222	Oxalobacteraceae family	-	-	-	✔

2.4.4 OTUs and fish fitness traits

Fish standard length, condition factor, splenosomatic or hepatosomatic indexes all significantly correlated with fish population (LM, Standard length: $F_{10,146}= 28.73$, $p < 0.001$, CF: $F_{10,146}= 6.84$, $p < 0.001$, SSI: $F_{10,144}= 18.01$, $p < 0.001$, HSI: $F_{10,146}= 5.18$, $p < 0.001$). We tested whether OTUs identified using ecotype pair SIMPERs were associated with fitness traits within ecotype pairs, to test for the possible influence of specific

microbes on host fitness. Specifically, the six detected OTUs in the Canadian ecotype pair were not associated with any difference in standard length, condition factor, splenosomatic or hepatosomatic indexes. In contrast, within the German ecotype pair 1 (Westensee – Eider) the presence of OTU 4342297, a *Clostridium* species (*Clostridium* species 1), correlated with fish body condition and splenosomatic index (LM, CF: $F_{1,25}=$ 7.12, $p < 0.05$ and SSI: $F_{1,25}= 6.77$, $p < 0.05$). *Clostridium* species 1 was found in greater abundance in the lake population than river population. The presence of *Clostridium* species 1 within Westensee correlated with increased body condition and a lower splenosomatic index (Figure 2.5A & B). A second *Clostridium* species (*Clostridium* species 2, OTU 73202) correlated with the standard length of fish within German ecotype pair 2 (Großer Plöner See – Malenter Au, LM, $F_{1,24}= 7.43$, $p < 0.05$). *Clostridium* species 2 was found in higher abundance in the river fish, and the presence of this bacterial species was negatively correlated with standard length (Figure 2.5C). OTU 222, a bacterial species belonging to the oxalobacteraceae family, was found in higher abundance in Malenter Au fish. OTU 222 was associated with changes in fish hepatosomatic index between populations from German ecotype pair 2 (LM, $F_{1,24}= 5.04$, $p < 0.05$, Figure 2.5D). A third OTU was associated with fish body condition and hepatosomatic index, *Propionibacterium acnes* (OTU 1087597), of German ecotype pair 2 fish (LM, CF: $F_{1,24}=$ 4.40, $p < 0.05$, Figure 4E and HSI: $F_{1,24}=5.35$, $p < 0.05$, Figure 2.5F). *Propionibacterium acnes* was found in higher abundance in the river population. For the Norwegian ecotype pair, the presence/absence of OTU 24, a bacterial species belonging to the brevinemataceae family was associated with differences in the body condition of fish (LM, $F_{1,36}= 5.23$, $p < 0.05$). The presence of OTU 24 was associated with increased body condition in fish from the Skogseidvatnet lake (Figure 2.5G). OTU 24 was found in higher abundance within the Norwegian ecotype pair in lake fish. The 18 Core OTUs identified were not correlated with any variation in fitness-related variables.
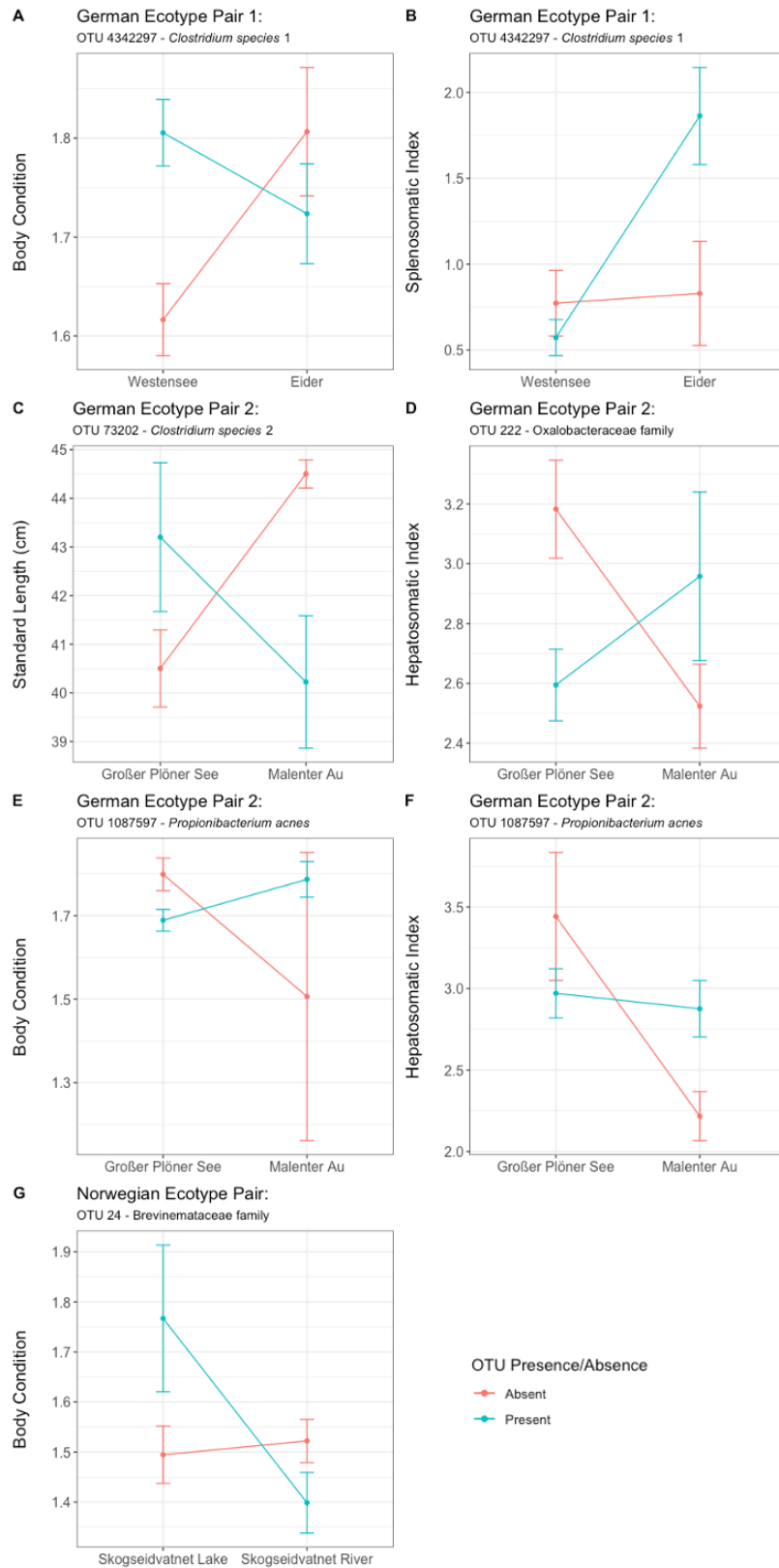
Figure 2.5: Reaction norms showing influence of presence/absence of bacterial species on different fish fitness traits: body condition, splenosomatic Index hepatosomatic index and standard length across three ecotype pairs: German ecotype pair 1: Westensee - Eider, German ecotype pair 2: Großer Plöner See – Malenter Au, a Norwegian ecotype pair: Skogseidvatnet lake - Skogseidvatnet river. Canadian ecotype pair not included as no fitness traits correlated with an interaction between OTU and fish population. Error bars represent standard error.

2.5 Discussion

Microbes affect their hosts' physiology as well as their life history traits and even their evolutionary trajectory. The three-spined stickleback is present in virtually all water bodies across the Northern hemisphere and expresses traits consistent with parallel and non-parallel evolution across their morphology and genomic architectures (Bolnick *et al.*, 2018, Feulner *et al.*, 2015, Hanson *et al.*, 2017). However, the factors contributing to the formation of their microbiomes remain elusive (Smith *et al.*, 2015). Here, we investigated whether fish carry a similar gut microbiome within habitat type (i.e. potential evidence for parallel evolution), or if microbiomes are best explained by population-specific communities (i.e. possible local adaptation). We found that bacterial diversity was best explained by population specificity. More importantly, we found that the main drivers of lake-river divergence were population pair specific and that microbial communities are best explained overall by fish population at both phylum and OTU levels. Among the 30 OTUs that varied in abundance within lake-river ecotype pairs, 5 were correlated with fish fitness. Altogether, the observed patterns suggest that local adaptation is a major determinant of fish microbiomes.

Gut microbes may facilitate colonisation of new habitats and improve the adaptive potential of species, through improved ability to process otherwise inadequate food sources, as well as playing a role in immune functions (Alberdi *et al.*, 2016, Suchodolski, 2011). We found 217 OTUs present in fish from our target populations, 93 of them were shared in fish across all ecotypes. The mixture of common and population-specific OTUs supports the theory that stickleback microbiomes are determined by a combination of environmental and genetic factors, as within a population fish have increased relatedness. Specifically, 17 core OTUs ($\geq 50\%$) were identified in fish from all populations suggesting these microbes are either specific to this species or a product of similar environmental factors across the different sample locations. The theory surrounding the

core microbiome also suggests that rather than OTUs being ubiquitous, it is the functions undertaken by the microbes that are core to the host-symbiont interaction (Huttenhower *et al.*, 2012). Yet, bacterial functionality currently remains unknown in sticklebacks so future research should explore the different functions specific bacteria may carry out within the stickleback microbiome. Phyla identified in this study are commonly found in the gut of other wild fish, although their roles within the environment or host remain to be elucidated (Llewellyn *et al.*, 2014, Smith *et al.*, 2015, Steury *et al.*, 2019).

Microbial diversity revealed differences at the phylum level among fish populations in a continent specific manner. Particularly, the North American river fish showed increased bacterial diversity compared to the other North American populations. In the European system, lake fish tended to show increased microbial diversity compared to river populations, with the highest diversity observed in German lake populations. This pattern was particularly visible at the phylum level, suggesting that while the lake river transition occurs in parallel across the northern continents, the underlying consequences on host-microbe interactions differ. Shannon's diversity at the phylum level also varied with sex, suggesting that sex-specific traits alter the microbiome. This relationship was strongest between females and males in North America, with females having higher alpha diversity. The stickleback sex genotype has previously been linked to the gut microbiome and revealed a genetic component mediating microbe diversity (Bolnick *et al.*, 2014c). Mammalian research has shown that variation in immune function and hormone production between the sexes can drive microbial variation (Markle *et al.*, 2013). Stickleback diet also varies with sex, and sex-specific infection due to trophically transmitted parasites (Brunner *et al.*, 2017, Eizaguirre *et al.*, 2009b), which may correlate with differences in microbial composition (Sanchez-Gonzales *et al.*, 2001). Furthermore, in larger fish we found that microbial diversity decreased. A potential explanation for this is that fish may become more fixed in their feeding ecology with age, reducing the

diversity of bacterial species a host of exposed to and revealing relationships between microbes and fish ontogeny (Sánchez-Gonzáles, Ruiz-Campos, & Contreras-Balderas, 2001). Overall, the combination of population-specific and sex-specific microbial diversity points towards a combination of environmental and genotypic factors driving an individual's gut microbiome diversity. The investigation of diversity at the phylum level enabled to detect patterns that were otherwise masked by the large variation observed at the OTU level.

Focusing on microbial communities allowed us to elucidate for the role of parallel evolution and local adaptation in host-symbiont interactions. Firstly, similarly to microbial diversity, we found that stickleback carried population-specific microbial communities at both the OTU and phylum levels. These population-specific communities translated into fish from different ecotypes and different continents having different microbial composition at OTU and phylum levels. Such a result suggests both that environmental determinism and evolutionary history contributes to gut communities. It is known that selection pressures across continents and ecotypes drive parallel and non-parallel evolution in stickleback morphology, behaviours, and population structure (Chain *et al.*, 2014, Feulner *et al.*, 2015). It is plausible that these selection pressures directly affect the community structure of the gut microbiome but also indirectly affect the host genetic architecture, which further impacts microbial composition. It is important to note that the observed patterns may also be driven by neutral processes and ecological flexibility of the microbiome, in order to confirm such findings evidence of genetic changes linked to the microbiome would be required.

Our findings suggest an important environmental role in shaping microbial communities; populations from the marine habitat showed consistent differences with fish from the freshwater system, while fewer differences existed within the freshwater system. Salt water is known to have a large effect on the gut microbiota of other fish such

as salmon (Dehler *et al.*, 2017, Steury *et al.*, 2019). The weak difference observed between lake and river ecotype microbiomes likely stem from the more similar abiotic factors occurring between freshwater systems, but also likely from the more recent colonisation of freshwater habitats resulting in reduced genetic divergence between freshwater ecotypes than to the ancestral marine ecotypes (Bell and Foster, 1994, Feulner *et al.*, 2015, Jones *et al.*, 2012). Both possibilities, with the environmental role being stronger, are supported by the observed isolation by distance with increasing microbe community differences with increasing distance (Bolnick *et al.*, 2014a, Bolnick *et al.*, 2014c, Smith *et al.*, 2015). Similarly, to microbial diversity, we also find correlations between microbial community structure and fish standard length. It is likely that the fish and the microbiota residing within them are adapting to the different environmental pressures that are unique to each population's local habitat.

Additionally, we find other lines of evidence that host-microbe interactions are population-specific. Firstly, within a habitat type, where selection pressures are the most similar, we found that bacterial communities were different, whether among lake populations or among river populations. Secondly, we find that microbial communities of each of the four ecotype pairs differed significantly to each other, suggesting ancestral host genetic background is important. Finally, within lake-river ecotype pairs, where lake and river populations are geographically connected and with related evolutionary histories, host populations harboured significantly different microbial communities. This pattern was detected for three of four lake-river transitions sampled in this study. All of our combined results demonstrate that host microbiomes are determined by factors specific to individual populations, which could be the product of localised environmental selective pressures and/or increased relatedness of fish within a population (genetic determination), resulting in population-specific communities. From a genetic point of view, these results are coherent with the genetic structure of most of these populations,

which show no shared genomic basis of adaptation across all lake or river ecotypes (Chain *et al.*, 2014, Feulner *et al.*, 2015, Lenz *et al.*, 2013, Marques *et al.*, 2016).

An ideal way to demonstrate whether the population-specific pattern of host microbial communities could lead to the evolution of local adaptation, and therefore reflects a strong link with a genetic determinism, is to show that hosts have increased fitness in the presence of specific microbes within a given lake-river ecotype pair. We find differences between fitness traits across fish populations, that might be driven by different selection pressures, such as fish environment or genetics, but it is possible that the presence or absence of certain microbiomes could also influence host fitness. Microbes can be beneficial by increasing the ability of nutrient uptake and improving host immune function (Lawley and Walker, 2013, Suchodolski, 2011). However, their beneficial presence can also be context dependent with some crossing the parasite-mutualist continuum, becoming pathogenic in the absence of more virulent pathogens or in stressful environments (Chamberlain *et al.*, 2014, King *et al.*, 2016). We found 5 OTUs correlated with fish fitness traits in a population-specific manner, with the presence of 3 bacterial species having negative effects within the population where they were found in a greater abundance. The presence of OTU 24, a bacterial species belonging to the brevinemataceae family, was correlated with increased fish body condition in Skogseidvatnet lake, where it was found in a greater abundance. The increased abundance of *Clostridium* species 1 (OTU 4342297) within Westensee was correlated in increased fish body condition and decreased splenosomatic index. Increased spleen size has been linked with costly immunological activation so a lower splenosomatic index and a higher body condition suggests fish are fitter (Kalbe *et al.*, 2009). On the other hand, the presence *Clostridium* species 2 within Malenter Au correlated with a reduction in standard length of fish, suggesting that the relationship between bacterial species and fish fitness is complicated, and closely related bacterial species may not affect fitness in the same way.

It may also be that host genetics between the populations lead to differences in host-microbe interactions. The *Clostridium* genus is known to contain species that are pathogenic to humans and animals (Gibbs, 2009, Turnbull *et al.*, 1990). While none of our results support a positive role of microbes in the colonization capacity of stickleback, they likely play an important role as agents of selection, with specific species acting differentially on different host genotypes (Ford *et al.*, 2017). It is important to note is that care must be taken in interpreting the association between fish fitness and the presence/absence of specific OTUs as these associations may arise by chance and other unmeasured factors may be the cause of the fitness variations. Overall, this analysis points towards a major environmental role, rather than a genetic role, in the determinism of gut microbes in natural stickleback populations.

While we identified population-specific microbial diversity and communities, quantifying the relative contribution of environmental and possible genetic factors on microbiome is challenging, even with replicated lake-river transitions. To this end, field experiments combining quantitative breeding and exposure to different environments should be conducted not only sampling microbiomes of fish but also that of food sources and water in order to observe the microbes present in the environment. Even though laboratory experiments may bring insights into host-symbiont interactions, they are unlikely to replicate the complexity of the natural environment and hence will remain inconclusive. Overall, our study shows host-symbiont interactions are influenced local environmental pressure, suggesting that local adaptation is a major determinant of stickleback gut microbiomes.

Chapter 3: Local adaptation shapes the joint evolution of host-parasite interactions and feeding ecology

Adrienne Kerley[1], Tina Henrich[2], Christoph L. Gahr[2], Martin Kalbe[2], Lee Henry[1], China Hanson[3], Manfred Milinski[2] and Christophe Eizaguirre[1]

[1]School of Biological and Chemical Sciences, Queen Mary University of London, London, United Kingdom.

[2]Max Planck Institute for Evolutionary Biology, Department of Evolutionary Ecology, August-Thienemann-Str. 2, 24306 Ploen, Germany.

[3] Microbiology@UCL, University College London, London WC1E 6BT, UK

Keywords:

Local adaptation, parasites, feeding ecology, stable isotopes, carbon, nitrogen

3.1 Abstract

Host-parasite and predator-prey interactions are strong drivers of local adaptation, but how they coevolve in a host population remains elusive. Here, we addressed this knowledge gap by performing a field-based common garden experiment using second-generation lake and river three-spined stickleback (*Gasterosteus aculeatus*) from Canada and Germany. Fish from both countries or ecotypes were placed into mesocosms located either in a lake or in a river in Germany, where they could prey on natural items and be infected by parasites naturally. After 10 months, fish were retrieved, and parasite load was estimated together with carbon (δ^{13}C) and nitrogen (δ^{15}N) stable isotope ratios as long-term proxy of feeding ecology. We found that fish exposed to lake conditions harboured more parasites than those exposed to river conditions. Interestingly, despite exposure to identical environmental conditions, parasite community were different in fish from Germany and Canada as well as between fish of different ecotypes (lake or river), suggesting different heritable resistance capacity among fish origins. Parasite load correlated with variation in the feeding ecology (δ^{13}C/δ^{15}N) and fish fitness proxies, and those correlations mostly revealed an advantage with respect to parasite resistance to the local fish over foreign fish. Together our results show that feeding ecology is not independent of parasite resistance and evolves under local adaptation.

3.2 Introduction

Local adaptation is a common outcome of species evolution (Kawecki and Ebert, 2004, Savolainen *et al.*, 2013, Sobel *et al.*, 2010). It is characterised by resident genotypes having higher relative fitness in their local habitat than non-residents (Kawecki and Ebert, 2004). Local adaptation is the result of the interaction between population evolutionary history and local environmental conditions, both abiotic and biotic (Kawecki and Ebert, 2004) with textbook examples often relating to host-parasite interactions (Berryman, 1992, Haldane, 1949, Summers *et al.*, 2003). Since multiple selection pressures occur at once in nature, host-parasite interactions are not independent of other aspects of a host's biology, such as feeding ecology (Anaya-Rojas *et al.*, 2016, Brunner *et al.*, 2017).

Parasites can drive the evolution of host behaviours, phenotypes, and even population genetic structure (Barber, 2013, Eizaguirre *et al.*, 2012b, Hamilton and Zuk, 1982, Milinski and Bakker, 1990, Møller, 1990). Indeed, differences in parasite abundance and diversity among host populations from contrasting habitat types can result in locally adapted immune genotypes, and, in specific cases, can also culminate into host speciation (Buckling and Rainey, 2002, Eizaguirre and Lenz, 2010, Eizaguirre *et al.*, 2012a, Eizaguirre *et al.*, 2009a).

Variation in feeding strategies can result in differences in exposure to both trophically transmitted and actively infecting parasites, through foraging in close proximity to infected individuals or in areas containing high numbers of free living parasites (Bakke *et al.*, 1992, Johnson *et al.*, 2009, Locke *et al.*, 2014, Marcogliese and Cone, 1997, Stutz *et al.*, 2014). For example, parasites that have intermediate invertebrate hosts that reside in oxygen rich littoral zones, infect more brook charr (*Salvelinus fontinalis*) feeding in benthic habitats than those that feeding in pelagic areas (Bertrand *et al.*, 2008). Additionally, generalist fish that feed on a large range of prey items are more likely to be exposed to and carry a more diverse community of parasite species (Locke *et*

al., 2014). Contrastingly, a specialist predator, predominantly feeding on a parasite's intermediate host would show an increased exposure to this parasite species (Marques *et al.*, 2011), likely mediating the evolution of a specific immune response, like that of river stickleback against *Gyrodactylus* sp. (Eizaguirre *et al.*, 2012a). As ontogenic development occurs, a host's diet can change and can result in diverse infection patterns within a population. For example, as yellow perch (*Perca flavescens*) age their feeding strategy shifts which correlates with increased helminth infections (Johnson *et al.*, 2004). Overall, those examples demonstrate how feeding and diets are associated with exposure to parasites.

Interestingly, individual differences in parasite load could also result in different feeding strategy (Brunner *et al.*, 2017). Specifically, dietary shifts after infection could be the result of parasite-mediated effects such as behavioural manipulations that impact feeding performance (Barber *et al.*, 2008, Lefevre *et al.*, 2009, Lochmiller and Deerenberg, 2000). Alternatively, a host may change feeding strategy, focusing on lower quality, easier to catch, food sources in order to still achieve a sufficient caloric intake (Milinski, 1984, Ponton *et al.*, 2011). Lastly, infected host may try to compensate the costs of infection from increased feeding of smaller less nutritious prey items (Brunner et al 2017). Hence, parasite infection both affects and is affected by host diet, highlighting the importance of understanding the combination of the evolution of resistance and feeding strategy.

Previous research on the relationship between feeding ecology and host-parasite driven local adaptation has mostly focused on gut content analysis (Bolnick *et al.*, 2020, Cirtwill *et al.*, 2016, Emde *et al.*, 2014, Kleinertz *et al.*, 2012, Reimchen and Nosil, 2001). Whilst gut content allows for the direct identification of food items, it (i) only reflects prey consumed shortly prior to sampling, (ii) can underestimate the amount of zooplankton consumed and (iii) if an individual's stomach is empty provides no

information of diet (Arrington *et al.*, 2002, Lafferty *et al.*, 2008, Matthews and Mazumder, 2005). Furthermore, items identified may not reflect the true nature of what is assimilated, as soft bodied items are digested first and leave no sign of past presence (Grey *et al.*, 2002). Stable isotope analysis on the other hand allows for a long-term assessment of an individual's diet, the timeframe of which depends on the tissue samples (Lorrain *et al.*, 2002, Post, 2002). Carbon isotope ratios (δ^{13}C) will generally be enriched by < 1 ‰ between a consumer and their food source. Such a change allows for the identification of the source of primary production in a food web and in turn enables making inferences of an organisms foraging habitat (DeNiro and Epstein, 1978, Fry, 2006). Similarly, the ratio of nitrogen isotope ratios (δ^{15}N) are enriched by 3-4 ‰ in consumers in comparison to their food source, indicating the trophic position of an organism (Minagawa and Wada, 1984, Post, 2002) and therefore its likely exposure to parasites with complex life cycles (Britton *et al.*, 2011, Pegg *et al.*, 2017).

The ecological interactions between parasite load and diet are difficult to replicate *in vitro* due to the complexity of both parasite and prey diversities. Therefore, *in situ* experiments are required to advance our understanding of how these ecological factors link to local adaptation. The three-spined stickleback (*Gasterosteus aculeatus*) is a good organism to test for host-parasite local adaptation and the influence of diet. Repeated colonisation of freshwater habitats across the northern hemisphere since the last glaciation has resulted in the evolution of multiple distinct ecotypes which differ in genetics, morphology, physiology, and behaviours (Bell and Foster, 1994, Eizaguirre *et al.*, 2011, Feulner *et al.*, 2015, McKinnon and Rundle, 2002, Rennison *et al.*, 2019b, Taylor and McPhail, 2000). Of great interest is the lake-river parapatric ecotype pair as their parasite community composition vary between lake and river ecotypes (Bolnick *et al.*, 2020, Eizaguirre *et al.*, 2011, Kalbe *et al.*, 2002, Reusch *et al.*, 2001). Specifically for the German system used in this study, lake fish are exposed to a greater diversity and load

of parasites than river fish (Eizaguirre *et al.*, 2011, Kalbe *et al.*, 2002), and therefore, lake ecotypes have evolved better resistance to a range of parasite taxa (Kalbe and Kurtz, 2006, Kurtz *et al.*, 2004). Furthermore, difference is resistance between ecotypes are associated with polymorphism at immune genes and their expression (Eizaguirre *et al.*, 2012a, Eizaguirre *et al.*, 2012b, Lenz *et al.*, 2013, Wegner *et al.*, 2003). Stickleback can be parasitised by a vast range of species, both actively and trophically (Barber, 2007, Barber, 2013, Stewart *et al.*, 2017), with virtually every part of their anatomy having the potential to be exploited by at least one parasite species (Kalbe *et al.*, 2002). Additionally, lake-river ecotype pairs show morphological differences and different feeding strategies (Berner *et al.*, 2008), which likely alter exposure to parasites (Stutz *et al.*, 2014). Lastly, variation in abiotic and biotic pressures among habitat types and the response of stickleback ecotypes to these heterogeneous habitats makes lake-river pairs a good natural system to explore local adaptation of geographically connected fish populations.

Here, we focused on two replicated lake-river ecotypes and hypothesised that if ecotype pairs from different systems show identical patterns of infection and diet preference, this would likely stem from parallel evolution. Alternatively, if we detect population-specific patterns, it suggests a relationship between infection (i.e., resistance) and diet evolved under local adaptation. To test those hypotheses, we conducted a field-based common garden experiment using lab-reared G2 three-spined stickleback from parapatric lake-river pairs from Germany and Canada, placed in an unfamiliar allopatric lake or river habitat in Germany. Fish were placed into allopatric habitats in Germany in order to allow for the accurate comparison of fish from different continents, as if German fish were placed into their sympatric habitat, they may have an additional advantage over Canadian fish. Fish were held in wire mesh mesocosms for 10 months. Individuals were screened for ecto and endo macroparasites. Fitness proxies were measured and stable

isotopes, $\delta^{13}C$ and $\delta^{15}N$, were assessed. This *in situ* experiment allowed us to explore the influence of parasites and diet without the constraints of a laboratory setting.

3.3 Methods

3.3.1 Breeding design

Wild three-spined sticklebacks were collected using a combination of hand netting and minnow traps from two connected lake and stream systems, one in Canada (McCreight lake: 50°28'12.4''N, 125°65'31.7''W, Amour de Cosmos creek: 50°23'54.3''N, 125°63'62.9''W) and one in Germany (Westensee lake: 54°26'89.8''N, 9°96'09.2''E, Eider stream: 54°16'65.5''N, 10°07'60.1''E). Individuals were collected during the 2014 and 2015 breeding season. First-generation (G1) families were bred from 20 individuals from each sampling population using *in vitro* fertilisation. After fertilisation, eggs were treated with acriflavine (Dajana) and Methylene Blue (King British) to prevent fungal infection. Fertilised Canadian eggs were transported to Germany at 4°C. Transportation conditions were replicated for German fish by storing fertilised eggs in the fridge at 4°C for 4 days. G1 individuals were then incubated, hatched, and raised in tanks in the laboratory with constant water flow and fed frozen Chironomidae sp. larvae *ad libitum* at 18°C, 18:6 Light:Dark (L:D). G1 fish were then cycled through different seasons artificially to initiate sexual maturity (Autumn: 12°C for 2 weeks, 12:12 L:D, Winter: 6°C for 4 weeks, 12:12 L:D, Spring: 12°C for 4 weeks, 12:12 L:D and Summer: 18°C until breeding, 18:6 L:D). Sexually mature males were kept in individual tanks and given nest materials, green polyester threads (cut to a length of ~ 10 cm) and a sand-filled petri dish. Unrelated gravid females from the same sampling population were placed into the tank with the males and allowed to spawn naturally. Breeding pairs were allowed to mate several times to make large family groups of identical genetic origin. We obtained a total of 24 second-generation (G2) families (N = 12 of Canadian origin and N = 12 of German origin).

3.3.2 Field-based common garden experiment

A dorsal spine clipping was taken for DNA extractions before fish were released into the experimental mesocosms. All fish were typed for 14 microsatellites using DNA Tissue kit (Invitek, Germany) following the manufacturer's protocol to allow for sex-typing and individual fingerprinting (Kalbe *et al.*, 2009). Length and weight of fish were measured at this stage. In October 2016, 24 experimental mesocosms, made of 5 mm stainless steel mesh (length: 1 m, height: 0.25 m, width: 0.6 m) were placed in two geographically connected habitats: Großer Plöner See a large lake (GPS, $N_{mesocosms} = 12$, $N_{fish} = 288$, 54˚14'61.0''N, 10˚40'86.9''E) and Malenter Au a small, slow-flowing stream (MAU, $N_{mesocosms} = 12$, $N_{fish} = 288$, 54˚19'62.7''N, 10˚55'65.9''E) following the protocol laid out in Eizaguirre *et al.* (2012a). The size of the metal mesh allowed for food, parasites, and water to flow freely through the mesocosms but kept experimental stickleback separate from wild stickleback populations and predators. These two sites were chosen to make sure that fish of German origin were exposed to a different lake-river system than their natal habitat, so all fish were in an allopatric system, allowing for more direct comparisons between fish of different country origins. Malenter Au mesocosms were placed in the centre of the stream bed 3 m apart at a depth of 0.5 – 1.5 m. Großer Plöner See mesocosms were placed 2 m apart at 1 – 1.2 m depth. Each mesocosm contained 12 river (♀ = 6, ♂ = 6) and 12 lake (♀ = 6, ♂ = 6) individuals from either Canadian or German origin. Sexes were separated by plexiglass down the centre of the mesocosm to prevent reproduction. Experimental mesocosms were visited weekly to check for dead fish and guarantee that mesocosms remained submerged, particularly in the stream habitat. Control mesocosms which were held in standardised, parasite free conditions in the laboratory ($N_{mesocosms} = 6$, $N_{fish} = 144$) were cycled through the seasons to mimic external environmental conditions. All fish were fed *ad libitum* on the same diet of frozen Chironomidae sp. larvae to measure stable isotope fractionation.

3.3.3 Sample Collection

After 10 months, fish were collected from the mesocosm, housed in aerated water from their habitat of exposure and dissected within two days of collection. The dissection period lasted 25 days overall. Fish were euthanised with tricaine methanesulfonate (MS222, 200mg/l, Sigma). External parts of fish were screened for ecto macroparasites under a dissection microscope. Eyes, gills and internal organs, except the spleen and the heart, were screened for endo macroparasites, using a compressorium (Kalbe *et al.*, 2002). Parasite abundance and diversity were combined into an individual parasite index, I_{PI} (Kalbe *et al.*, 2002). Standard length (SL), total length (TL), weight, and organ weight (liver, gonads, spleen and head kidney) were measured. Two fitness measures were created for body condition (CF $= \left(\frac{Body\ Weight}{Standard\ Length}\right)^3 \times 100$) (Frischknecht, 1993) and splenosomatic index (SSI $= \frac{Spleen\ Weight}{Body\ Weight} \times 100$) (Kalbe *et al.*, 2009). DNA samples were taken to identify individuals, their family background and to quantify their growth rate. A small section (~ 0.5 cm) of white muscle was dissected from the tail and frozen at -20°C for stable isotope analysis.

3.3.3.1 Stable Isotope analysis

Muscle samples were dried at 60°C for 48 hours and homogenised using a mechanical grinder for 1 minute (30Hz, 2 x 316 stainless steel beads, TissueLyser II, Qiagen). Tin capsules (6 x 4 mm). were filled with approximately 1 mg of homogenised sample material and folded shut. Samples were analysed on an Integra 2 spectrometer (Sercon Instruments, Crewer, UK, Analytical precision: 0.1‰) using continuous flow isotope ratio mass spectrometry (CF-IRMS). Carbon (δ^{13}C) and nitrogen (δ^{15}N) isotope ratios were calculated as proxies for fish feeding ecology. Accuracy and precision of the sample runs was verified every 10 samples using a Protein (Casein) Standard

OAS/Isotope to check deviation of $\delta^{13}C$ and $\delta^{15}N$ readings (Elemental Microanalysis, $\delta^{13}C$: -26.98 ± 0.13‰, $\delta^{15}N$: +5.94 ± 0.08‰).

3.3.4 Data analyses

Statistical analyses were conducted in R statistical package version 3.6.3 (R Core Team, 2020). R packages *lme4* and *lmerTest* were used for fitting linear mixed effect models (LMMs). We tested the normality of model residuals and data transformation was conducted if required to meet test assumptions. All models were backwards selected using Akaike's Information Criterion (AIC) values to retain the optimal reduced model (Burnham and Anderson, 2004). Collinearity between fixed variables within a model were tested and if observed, residuals of their regressions were used.

3.3.4.1 Testing for fish survival differential

The ability for an organism to survive in an environment is key for the evolution of local adaptation. To test survival rates of fish, we used a binomial generalised linear mixed effect model (GLMM) with fish country of origin (Canada or Germany), habitat of exposure (Großer Plöner See or Malenter Au or Laboratory) ecotype (lake or river) as well as their interaction as fixed variables. Mesocosm ID, and sex were set as random effects. Family group could not be included as a random effect as the family group of dead fish could not be confirmed by molecular sampling. Pairwise comparisons were conducted between categories of the significant variables using Tukey HSD post-hoc tests.

3.3.4.2 Stable isotope fractionation

Firstly, to understand the possible metabolic differences between fish from different countries and ecotypes, we tested for fractionation difference of stable isotopes. Laboratory fish, which were kept in standardised housing and feeding conditions, in mixed families, were used in two separate linear mixed effects models, for $\delta^{13}C$ and $\delta^{15}N$,

with fish country of origin, ecotype and their interaction as fixed factors, with fish family, sex and mesocosm ID as random factors.

3.3.4.3 Determinants of fish feeding ecology

To identify the determinants of fish feeding ecology, we performed a series of linear mixed effect models for $\delta^{13}C$ and $\delta^{15}N$ independently. Fish country of origin, ecotype, habitat of exposure, parasite load (I_{PI}) and their interactions were used as independent variables. Family, sex and mesocosm ID were set as random factors to account for any family-based genetic variation, sexual-dimorphism or possible bias driven by mesocosm positioning in the different habitats. Parasite load was collinear with habitat of exposure and ecotype therefore the residuals of a linear model including those variables were used.

We then split the dataset into habitat of exposure, Großer Plöner See and Malenter Au to focus on the local vs foreign theory of local adaptation and test whether foreign ecotypes or countries of origin were correlated with changes in feeding ecology within the same environmental conditions. For each exposure habitat, we ran two linear models, $\delta^{13}C$ and $\delta^{15}N$, with country of origin, ecotype and parasite load (expressed as residuals of the regression between parasite load and ecotype) as fixed factors and family, sex and mesocosm ID as random factors.

To understand whether the presence of an individual parasite taxon was associated with different feeding ecologies, we performed two linear mixed effects models. To retain statistical power, we focused on the eye fluke, *Diplostomum* sp. and the flatworm, *Gyrodactylus* sp. as both parasite taxa were found with >10 % prevalence in both habitats of exposure. Two LMMs were used to test for correlations between $\delta^{13}C$ or $\delta^{15}N$ and fish habitat of exposure, country of origin, fish ecotype, parasite taxa abundance and their interactions. Family, sex and mesocosm ID set as random factors. *Diplostomum* sp. abundance was collinear with habitat of exposures and ecotype so the residuals of the

regression between *Diplostomum* sp. abundance, habitat of exposure and ecotype were used. *Gyrodactylus* sp. abundance was also collinear with ecotype so residuals of the regression between *Gyrodactylus* sp. abundance and ecotype were used in the models including *Gyrodactylus*.

3.3.4.4 Estimating plasticity in fish feeding ecology

In order to survive in a new or changing environment an organism must adjust often relying on adaptive phenotypic plasticity to do so (Crispo, 2007, Fitzpatrick, 2012). Our experimental design allows to investigate the plastic potential of fish as individuals are placed in either a similar habitat to which they originated from (native) or into an alternative habitat. To test how ecotypes adjust to novel prey and become infected in the reciprocal habitat, we created a variable that quantifies isotopic change between native and reciprocal habitats of fish families that are of lake or river origin (referred to as "feeding adjustment factor"). The capacity for a family group to adjust to their novel reciprocal habitat was calculated using equation 1 for lake families and equation 2 for river families. For example, fish from lake families exposed to river conditions (MAU) show a feeding adjustment factor of 0, this means they exploit the same niche as the river fish within the river habitat. This would show strong phenotypic plasticity of lake fish. If this index deviates significantly from 0, this suggests constraints into using the new niche, as well as weaker phenotypic plasticity and the use of sub-optimal feeding ecology. A linear model (LM) was used to test if feeding adjustment factor, and therefore feeding plasticity, differed with the country of origin, fish ecotype and their interaction.

$$A\ lake\ families\ adjustment\ factor: \frac{(Mean\ \delta_{Family\ in\ lake} - Mean\ \delta_{Family\ in\ river})}{(Mean\ \delta_{Family\ in\ lake} - Mean\ \delta_{All\ river\ ecotype\ families\ in\ river})} \quad \dots Eq.\ 1$$

$$A\ river\ families\ adjustment\ factor: \frac{(Mean\ \delta_{Family\ in\ river} - Mean\ \delta_{Family\ in\ lake})}{(Mean\ \delta_{Family\ in\ river} - Mean\ \delta_{All\ lake\ ecotype\ families\ in\ lake})} \quad \dots Eq.\ 2$$

3.3.4.5 Parasite load and community structure

Once the links between stable isotope ratios and parasite load had been identified, we wanted to understand the determinants of parasite load (expressed as individual parasite load, I_{PI}, Kalbe *et al.*, 2002). To this end, we performed a LMM with habitats of exposure, country of origin, ecotypes, sex and their interactions as explanatory variables with family and mesocosm ID as random factors. Within a habitat of exposure, similar parasite loads across both fish ecotypes and origins would suggest parasite load is solely determined by a host's environment. On the other hand, similar parasite loads within an ecotype but differences between ecotypes would suggest parallel ecotype evolution of parasite resistance. Lastly, population-specific parasite load within habitat of exposure would be evidence for local adaptation to be the main driver of parasite resistance/susceptibility. We used Tukey's honest significance post-hoc tests (Tukey HSD) for pairwise comparisons of interactions.

To investigate the determinant of parasite communities, parasite abundances were square-root transformed before performing a PERMANOVA with Bray-Curtis dissimilarity matrix. We used habitat of exposure, country of origin, fish ecotypes and their interactions as explanatory variables. Significant effects were followed by a dispersion test with *betadisper*, from the vegan package, to determine if the observed differences were a by-product of data dispersion.

After this initial test, fish were then split by habitat of exposure to further explore difference in parasite community, as parasite diversity and abundance vary greatly between habitat types (Eizaguirre *et al.*, 2011, Kalbe *et al.*, 2002). The correlation between country of origin, fish ecotype and their interaction and habitat of exposure specific parasite communities were tested using a PERMANOVA. Pairwise PERMANOVAs between significant factors were conducted using the *pairwise.adonis2* function from the pairwiseAdonis package. A dispersion test was carried out on all

significant effects. Distance-based redundancy analysis (dbRDA) were used to visualise and test the influence of $\delta^{13}C$ and $\delta^{15}N$ values on lake exposure and river exposed parasite communities separately.

3.3.4.6 Parasite load and fish fitness

In order to evaluate whether parasite load impacted fish fitness proxies, we conducted series of a linear mixed effects models on fish growth rate, both in length (length growth rate) and in weight (weight growth rate), body condition (condition faction, CF) and splenosomatic index (SSI). CF is a ratio between observed weight and the expected weight estimates from the observed length. Additionally, it has been well characterised that an increased SSI is associated with costly immunological cell activation and reduced life time reproductive success in this fish species (Kalbe *et al.*, 2009). For each fitness proxy, we tested the effect of parasite load, fish country of origin, habitat of exposure, and fish ecotype, as well as their interactions using a LMM. Family, sex and Mesocosm ID were set as random factors.

3.4 Results

High lipid content can bias stable isotope analysis considerably, altering the $\delta^{13}C$ value as lipids are depleted in ^{13}C, calling for possible mathematical lipid correction if samples have greater than 5% lipid content (Carbon to nitrogen ratio, C:N ratio > 3.5, Post *et al.*, 2007). In this study, lipid correction was not necessary as fish C:N ratios indicated low lipid content (C:N ratio: 3.3 ± 0.1).

3.4.1 Fish survival differential

A total of 459 fish survived the 10 months of the experimental period (N_{total} = 720). We found the probability of a fish surviving correlated with an interaction between their country of origin and habitat of exposure (GLMM, $X^2_2 = 22.85$, $p < 0.001$). Canadian fish broadly had a lower survival rate than German fish in experimental habitats, for example only 12.5% of the Canadian fish placed into Großer Plöner See survive the entire

course of the 10-month experiment (Table 3.1). Only Canadian fish from Großer Plöner See had a significantly increased mortality rate compared to German fish across all habitats of exposure (TukeyHSD, p < 0.001, Supplementary Table 3.1). Fish ecotype did not correlate with survival rate (GLMM, $X^2_1 = 0.27$, p = 0.6).

Table 3.1: Total survival rate of fish across different habitats of exposure and country of origin. Großer Plöner See – lake treatments, Malenter Au – river treatment, Lab – Laboratory control treatment

Habitat of Exposure	Country of Origin	Total Survival (%)
Großer Plöner See	Canada	12.5
	Germany	78.08
Malenter Au	Canada	66.9
	Germany	78.08
Lab	Canada	94.59
	Germany	81.94

3.4.2 Feeding ecology

3.4.2.1 Evaluating stable isotope fractionation

How an organism assimilates the organic material they consume is important but different fractionation rates of stable isotopes can occur. Here, we compared how fish from different origins held under standardised feeding conditions in the laboratory assimilated food. Under controlled conditions, river fish ecotypes showed higher $\delta^{13}C$ than lake fish (LMM, $F_{1,22} = 8.53$, p < 0.01, river fish: -18.66 ± 0.51, mean ± standard deviation, SD, lake fish: -18.98 ± 0.49). Fish country of origin was not associated with any differences in carbon fractionation (LMM, $F_{1,7} = 3.66$, p = 0.099). Neither country of origin nor ecotype was correlated with $\delta^{15}N$ values of laboratory fish (LMM, Country: $F_{1,5} = 1.99$, p = 0.218, Ecotype: $F_{1,20} = 2.18$, p = 0.155). This result suggests there could be some genetic differences in fractionation between fish ecotype and therefore call for caution when interpreting patterns observed for $\delta^{13}C$ variation.

3.4.2.2 Determinants of fish feeding ecology

To explore the complex relationship between host-parasite interactions and host diet in the wild, we used fish that had been held in the experimental mesocosms under natural environments. Fish δ^{13}C correlated with a four-way interaction between the habitat of exposure, country of origin, ecotype, and parasite load (LMM, $F_{1,310} = 5.03$, p < 0.05). δ^{15}N values were associated with a three-way interaction between the habitat of exposure, country of origin and ecotype (LMM, $F_{1,320} = 10.95$, p < 0.01). Parasite load was removed from the model through backwards selection.

Due to the complex nature of the four-way interaction, we split the dataset by habitat of exposure and ran LMMs on Großer Plöner See and Malenter Au separately. This approach further allowed us to focus on the local vs. foreign theory of local adaptation.

The δ^{13}C value for Großer Plöner See fish correlated with an interaction between fish country of origin and parasite load (LMM, $F_{1,107} = 4.62$, p < 0.05, Figure 3.1). Fish from both countries of origin showed a positive correlation between δ^{13}C values and parasite load, but the correlation was stronger in Canadian fish (Figure 3.1). No variables tested were significantly correlated with δ^{15}N values of fish held in the Großer Plöner See lake, suggesting regardless of country of origin, ecotype and parasite load, fish feed at similar trophic levels (LMM, country of origin: $F_{1,22} = 0.95$, p $= 0.339$, ecotype: $F_{1,24} = 0.63$, p $= 0.436$ and parasite load: $F_{1,24} = 0.63$, p $= 0.436$).

Within the Malenter Au habitat, we found an interaction between fish ecotype and parasite load correlated with $\delta^{13}C$ values, whereby lake fish showed a positive correlation between $\delta^{13}C$ values and parasite load, while this correlation was negative for river fish (LMM, $F_{1,197} = 4.97$, $p < 0.05$, Figure 3.2). The increase of $\delta^{13}C$ values in lake fish, a foreign ecotype, as they are more heavily parasitised suggests a dietary shift to more benthic food sources, which is not observed in river fish, the local ecotype. No correlations were observed between $\delta^{15}N$ values and fish country of origin, ecotype or parasite load (LMM, country of origin: $F_{1,13} = 3.21$, $p = 0.096$, ecotype: $F_{1,22} = 1.59$, $p = 0.219$ and parasite load: $F_{1,197} = 0.4$, $p = 0.528$).
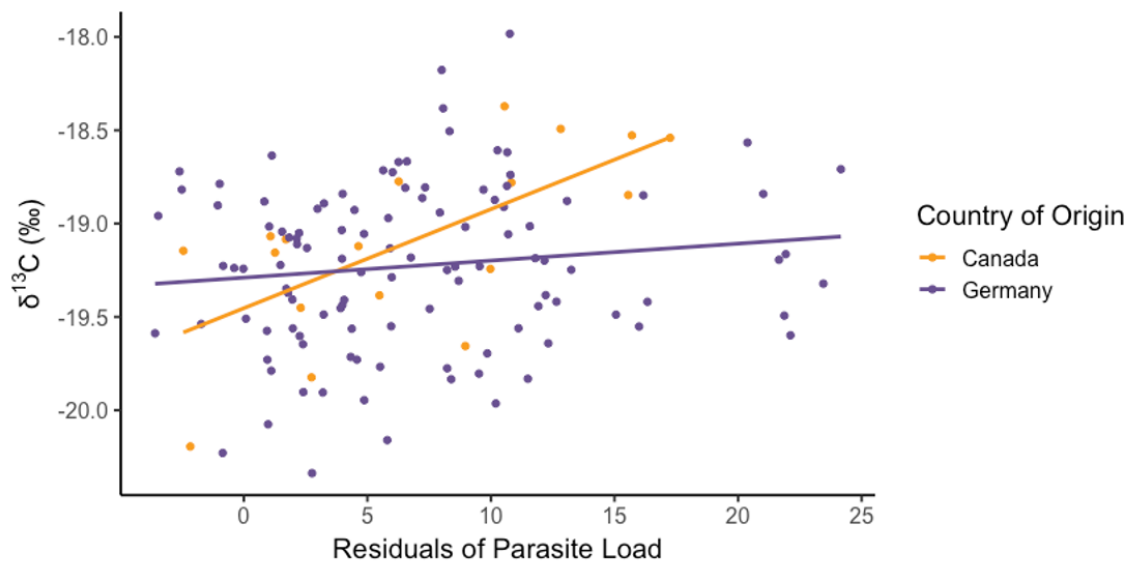
Figure 3.1: $\delta^{13}C$ of fish from Großer Plöner See correlated with residuals of fish parasite load (I_{PI}) and country of origin (LMM, $F_{1,107} = 4.62$, $p < 0.05$).
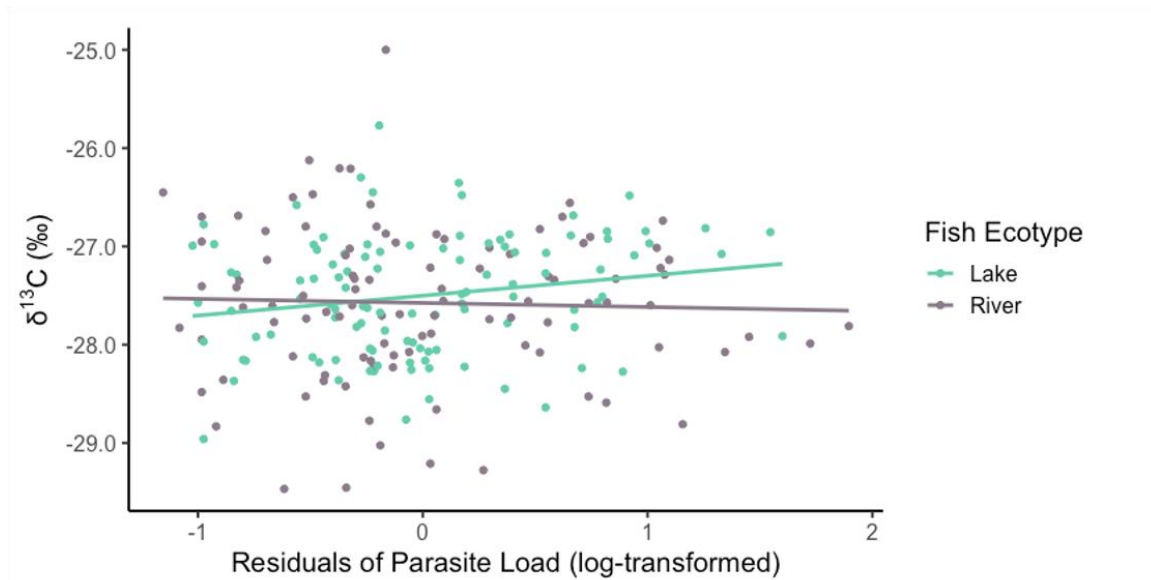
Figure 3.2: $\delta^{13}C$ of fish from Malenter Au correlated with residuals of fish parasite load (I_{PI}) and ecotype (LMM, F1,197 = 4.97, p < 0.05). Residuals of parasite load log-transformed for visual representation.

3.4.3 The influence of specific parasite taxa on fish feeding ecology

Specific parasite taxa can have a huge impact on a host's fitness and behaviour such as the feeding ecology of their host and how feeding varies among fish from different origins and within different habitats of exposure. We therefore, focused on *Diplostomum* sp. and *Gyrodactylus* sp. as they were found in both habitats with a prevalence >10 % (Supplementary Table 3.2), with *Diplostomum* sp. being a common lake parasite and *Gyrodactylus* sp., the main species in the river habitat (Eizaguirre *et al.*, 2012a) *Diplostomum* sp. did not correlate with variation in $\delta^{13}C$ values and was removed from the final model. However, fish $\delta^{15}N$ positively correlated with *Diplostomum* sp. abundance (LMM, $F_{1,215}$ = 8.79, p < 0.01, Figure 3.3) suggesting more heavily infected individuals feed at a higher trophic position, independently of the habitat of exposure.

Fish $\delta^{13}C$ values correlated with an interaction between *Gyrodactylus* sp. abundance and fish ecotype: $\delta^{13}C$ values of lake fish were negatively correlated with *Gyrodactylus* abundance and, whilst river ecotypes showed a positive correlation (LMM, $F_{1,316}$ = 5.87, p < 0.05, Figure 3.4). *Gyrodactylus* sp. abundance was not correlated with $\delta^{15}N$ values and was removed from the model through backwards selection.
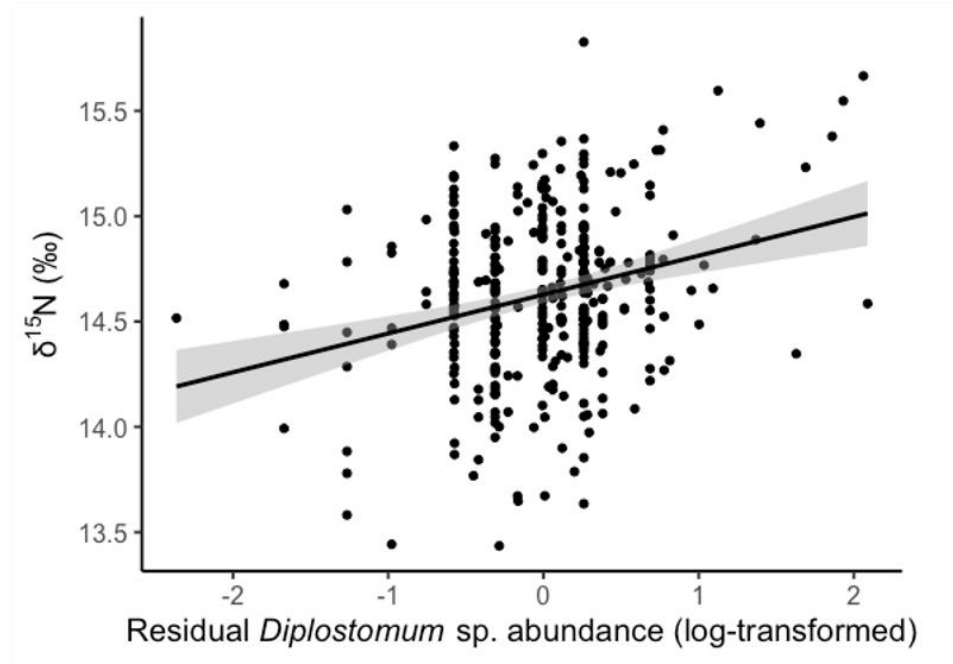
Figure 3.3: Fish $\delta^{15}N$ value was positively correlated with residuals of Diplostomum sp. abundance – residuals used (LMM, $F_{1,215} = 8.79$, $p < 0.01$). Residuals of Diplostomum sp. abundance log-transformed for visual representation.
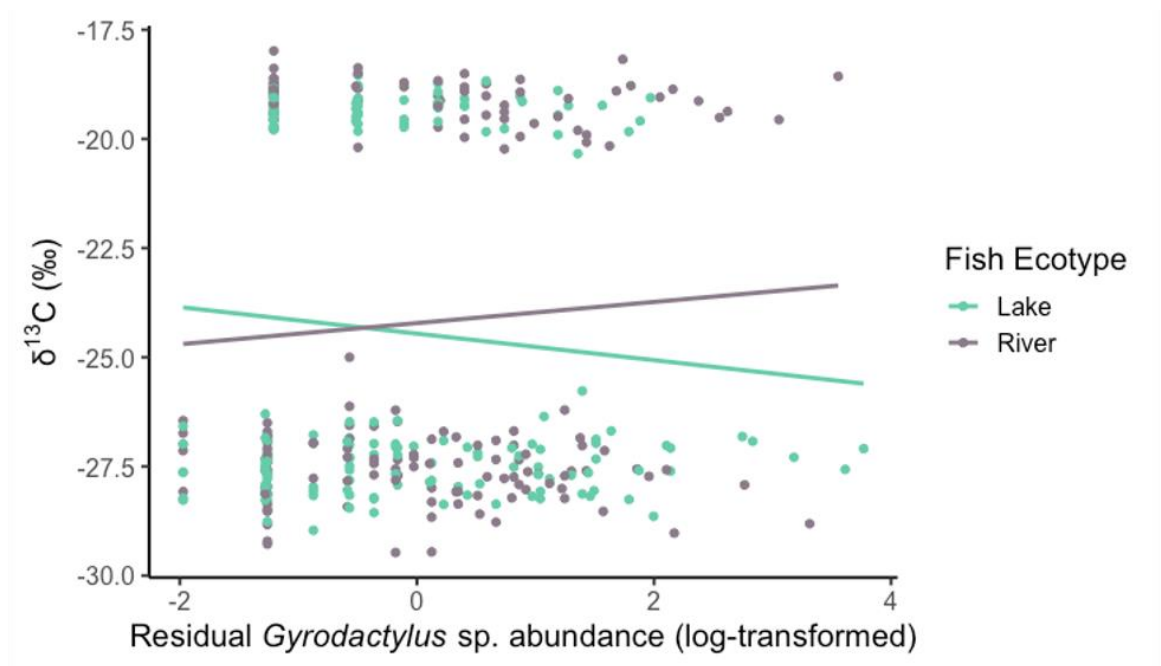
Figure 3.4: Fish $\delta^{13}C$ values correlated with an interaction between residuals of Gyrodactylus abundance and fish ecotype (LMM, $F_{1,316} = 5.87$, $p < 0.05$). Residuals of Gyrodactylus abundance log-transformed for visual representation

3.4.3.1 Estimating plasticity in feeding ecology

Because there were marked differences between fish ecotypes, but also between their country of origin and habitat of exposure, we created a feeding adjustment factor for $\delta^{13}C$ and $\delta^{15}N$ to determine the plastic ability of the lake fish to exploit a similar niche as the river fish in the river habitat, and vice versa. Lake ecotypes showed significantly greater variation in their $\delta^{13}C$ feeding adjustment factor than river ecotypes (lake ecotype: mean $\pm$ SD, 1.10 ± 0.02, river ecotype: 1.00 ± 0.05, LM, $F_{1,20} = 35.67$, $p < 0.001$, Figure 3.5A) suggesting that lake families were more capable of changing their feeding behaviour to match that of river fish, suggesting a higher plastic ability. Country of origin did not correlate with the $\delta^{13}C$ feeding adjustment factor and was dropped from the model during the model selection process. Neither country of origin nor ecotype correlated with the $\delta^{15}N$ feeding adjustment factor (LM, country: $F_{1,18} = 2.18$, $p = 0.157$ and ecotype: $F_{1,18} = 0.02$, $p = 0.877$, Figure 3.5B). Of note, even with the two extreme $\delta^{15}N$ feeding adjustment factor outliers (family: 123x608 and 114x622) removed, neither country of origin nor ecotype are significantly correlated with the $\delta^{15}N$ feeding adjustment factor.
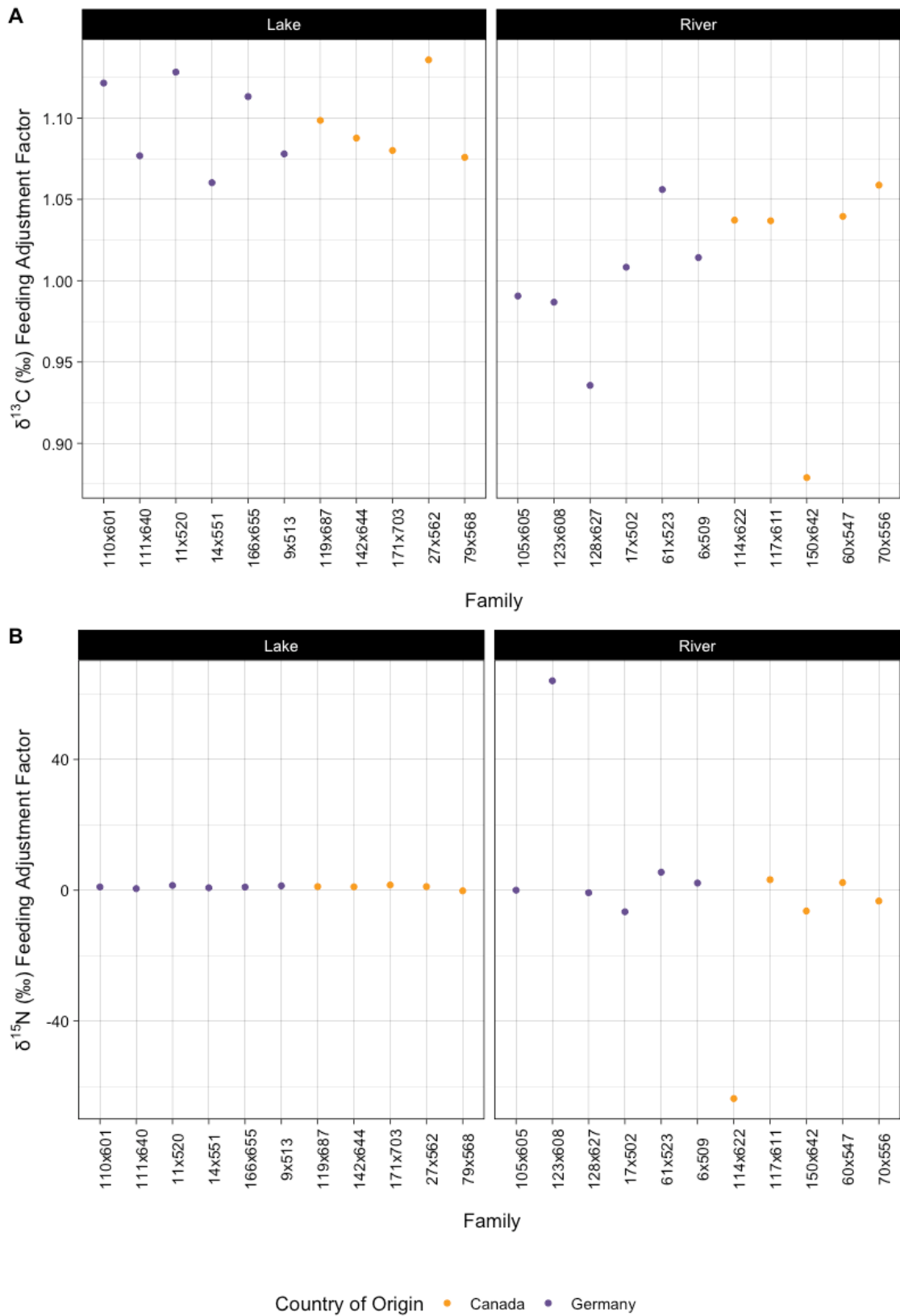
Figure 3.5: Plots showing A) $\delta^{13}C$ feeding adjustment factor and B) $\delta^{15}N$ feeding adjustment factor by family, ecotype and country of origin.

3.4.4 Parasite load and community structure

Since we determined that stable isotope ratios of $\delta^{13}C$ and $\delta^{15}N$ correlate with the parasitic environment the fish were exposed to, we undertook to understand the determinant of parasite load and communities in our experimental fish. A total of 19 parasite species were identified. Six taxa were present in both habitats of exposure, 11 parasite species were specific to Großer Plöner See lake and two species, Apiosoma sp. and Acanthocephalus lucii, were found only in the Malenter Au river (Supplementary Table 3.2). Parasite load correlated with a three-way interaction including the fish habitat of exposure, their country of origin and ecotype (LMM, $F_{1,323} = 5.23$, $p < 0.05$, Figure 3.6A, Supplementary Table 3.3). Post-hoc tests revealed fish from different habitats of exposure had different parasite loads; for example, Canadian lake fish in Großer Plöner See had a higher parasite load than Canadian lake fish in Malenter Au. Such a significant comparison is not surprising as there are more parasites in lake habitats than river (Tukey HSD, $p < 0.05$, Supplementary Table 3.4). However, no significant pairwise differences were found within each habitat of exposure (Tukey HSD, $p > 0.05$, Supplementary Table 3.4). Nonetheless, we found an ecotype by habitat of exposure effect, with river fish in Großer Plöner See being more heavily parasitised than lake fish, whilst within Malenter Au no significant difference in parasite load occurs (LMM, $F_{1,323} = 4.73$, $p < 0.05$, Figure 3.6B, Supplementary Table 3.3, Supplementary Table 3.5). We also found that males had a lower parasite load than females overall (LMM, $F_{1,316} = 5.22$, $p < 0.05$, Supplementary Table 3.3).
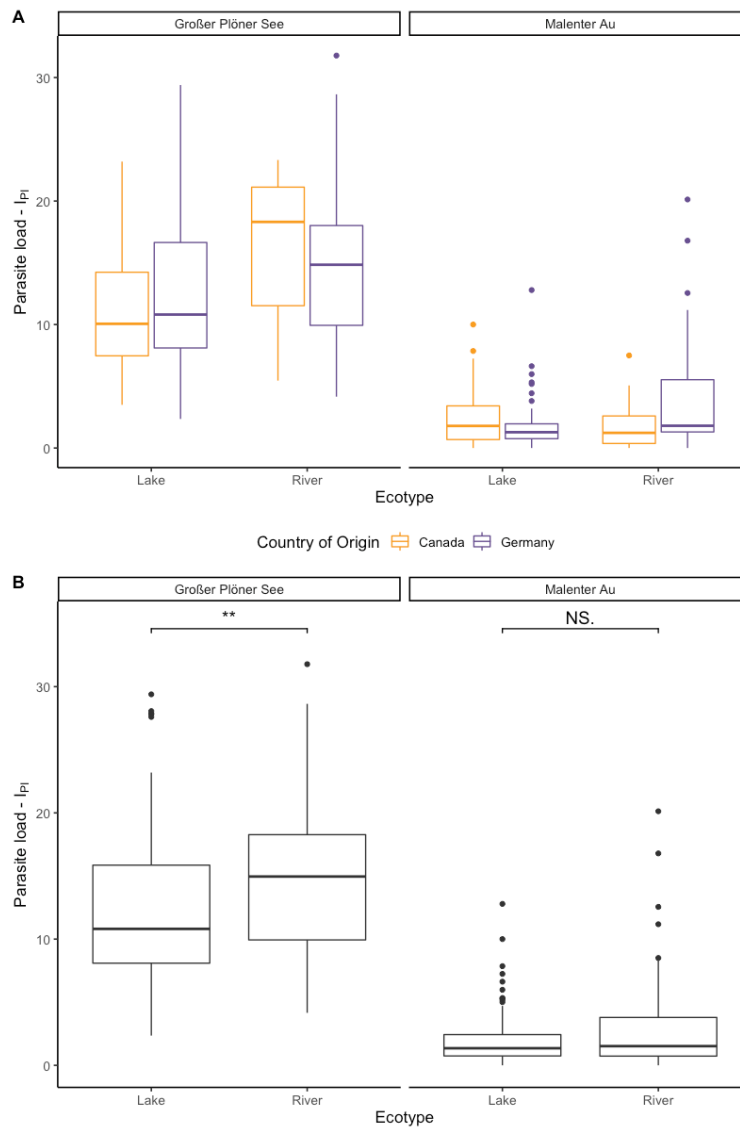
Figure 3.6: Relationship between parasite load (I_{PI}) and A) fish habitat of exposure, country of origin and ecotype, (LMM, $F1,323 = 5.23$, $p < 0.05$) and B) fish habitat of exposure and ecotype (LMM, $F1,323= 4.73$, $p < 0.05$).

When investigating the determinants of parasite community, we identified a significant interaction between fish habitat of exposure, country of origin and ecotype (PERMANOVA, $F_{1,337}= 2.49$, $p < 0.05$). A pairwise PERMANOVA showed that all pairwise comparisons between the two habitats of exposure were significant (pairwise PERMANOVA, $p < 0.05$, Supplementary Table 3.6). Within Großer Plöner See, German lake fish harboured significantly different parasite communities in comparison to German river fish and either Canadian ecotype, whilst in Malenter Au only Canada river fish harboured significantly different parasite to German river fish (pairwise PERMANOVA,

p < 0.05, Supplementary Table 3.6). Noteworthy, heterogenous data dispersion may contribute to this pattern (betadisper, $F_{1,330}$= 4.80, p < 0.001).

In order to explore this relationship further, we split fish by their habitats of exposure as parasite loads vary greatly between lake and river habitat. Parasite community differed significantly within Großer Plöner See with a country of origin by ecotype effect, (PERMANOVA, $F_{1,129}$= 2.53, p < 0.05, Figure 3.7A). Canadian lake and river ecotypes had significantly different parasite communities to German lake and river ecotypes within the lake (pairwise PERMANOVA, p < 0.05, Supplementary Table 3.7). Additionally, German river fish ecotypes harboured significantly different communities than German lake fish ecotypes (pairwise PERMANOVA, p < 0.05, Supplementary Table 3.7). Here as well, those heterogeneous groups showed significant dispersion of data (betadisper, $F_{1,126}$= 8.56, p < 0.001).

Similarly to Großer Plöner See, parasite community of Malenter Au fish, correlated with a country of origin by ecotype effect (PERMANOVA, $F_{1,207}$= 4.85, p < 0.01, Figure 3.7B). However, in the river habitat only one pairwise analysis revealed significant differences, involving the Canadian river ecotype and German river ecotype (pairwise PERMANOVA, p < 0.05, Supplementary Table 3.7). Here as well data dispersion may impact the observed results (betadisper, $F_{1,204}$= 2.97, p < 0.03).

Overall, clear evidence of differences is seen in the lake habitats whereby the parasite community is diverse, enabling the genetic effects associated with ecotype and country of origin to be detected (Figure 3.7).
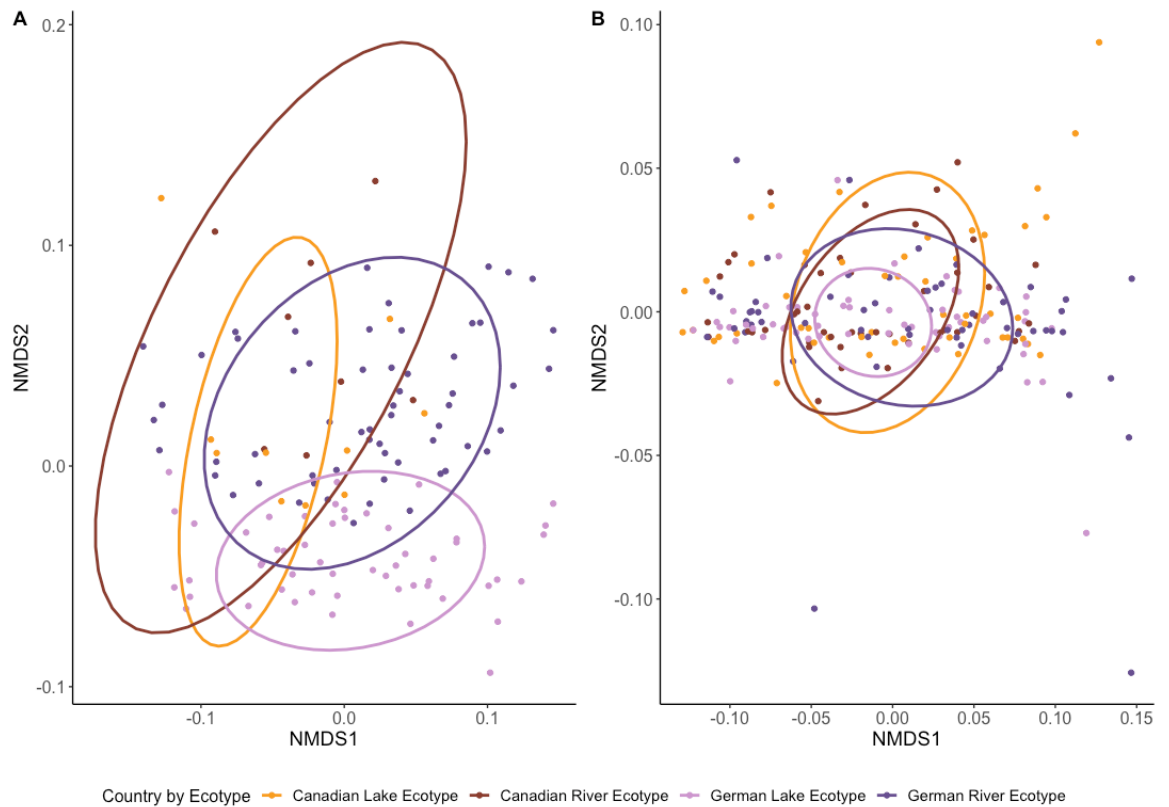
Figure 3.7: Non-metric multidimensional scaling (NMDS) plot of A) the parasite community of fish held within Großer Plöner See based on Bray-Curtis distance coloured by fish country of origin and ecotype (PERMANOVA, $F_{1,129}= 2.53$, $p < 0.05$) and B) the parasite community of fish held within Malenter Au based on Bray-Curtis distance coloured by fish country of origin and ecotype (PERMANOVA, $F_{1,207}= 4.85$, $p < 0.01$).

Distance-Based Redundancy Analysis (dbRDA) allowed us to correlate parasite communities with feeding ecology as shown by stable isotope values. Within the Großer Plöner See lake, $\delta^{13}C$ and $\delta^{15}N$ values were correlated with different parasite community composition (dbRDA, $\delta^{13}C$: $F_{1,127} = 3.14$, $p < 0.01$, $\delta^{15}N$: $F_{1,127} = 5.88$, $p < 0.001$, Figure 3.8A). Similarly, parasite communities of fish exposed to the Malenter Au river, correlated with $\delta^{15}N$ values but not $\delta^{13}C$ (dbRDA, $\delta^{15}N$: $F_{1,205} = 4.65$, $p < 0.01$, $\delta^{13}C$: $F_{1,205} = 0.61$, $p = 0.61$, Figure 3.8B).

103
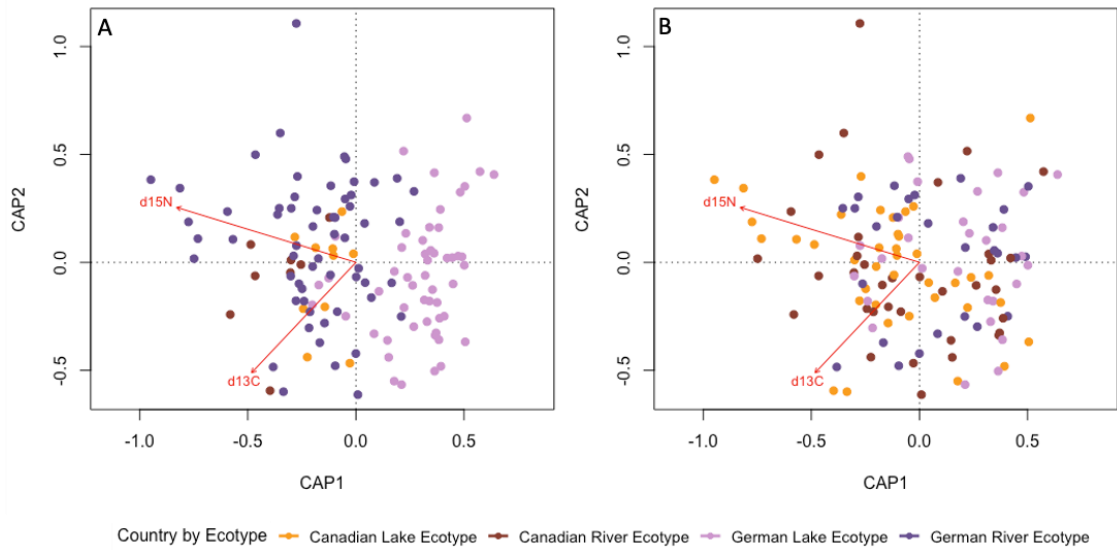
Figure 3.8: dbRDA plots showing Bray Curtis matrix on parasite communities with arrows representing δ¹³C and δ¹⁵N in A) Großer Plöner See (dbRDA, δ13C: F1,127 = 3.14, p = 0.002, δ¹⁵N: F1,127 = 5.88, p = 0.001) and B) Malenter Au (dbRDA, δ¹⁵N: F1,205 = 4.65, p = 0.006).

3.4.5 Parasite load and fish fitness

Because parasites impose a cost to their hosts, we tested how parasite load correlated with fitness proxies. Fish growth rate (calculated from length) correlated with an interaction between parasite load and country of origin, whereby Canadian fish growth rate correlated positively with parasite load, whilst German fish showed a weaker positive correlation (LMM, $F_{1,310}$= 5.44, p < 0.05, Figure 3.9, Supplementary Table 3.8). Fish SSI correlated with an interaction between parasite load, habitat of exposure and ecotype: lake fish ecotypes in the Großer Plöner See lake showed weak negative correlation between SSI and parasite load, whilst the parasite load of river and lake ecotypes held in the Malenter Au river positively correlated with SSI (LMM, $F_{1,322}$ = 10.45, p < 0.01, Figure 3.10, Supplementary Table 3.8). Parasite load did not correlate with fish weight growth rate or CF.
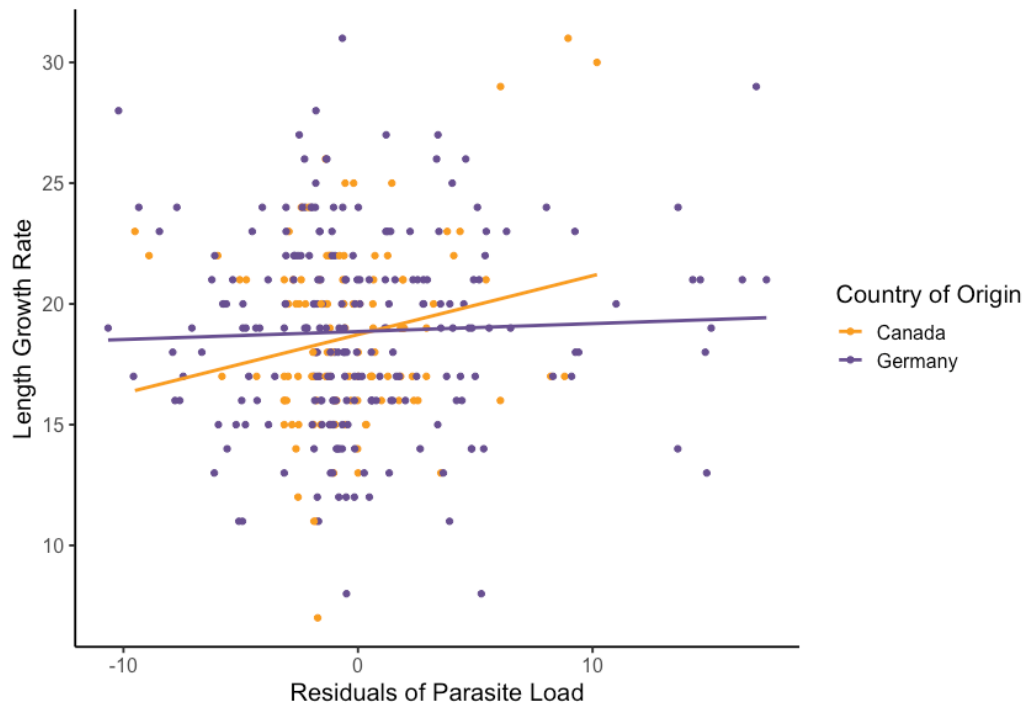
Figure 3.9: Fish length growth rate correlated with an interaction between parasite load and country of origin, (LMM, $F_{1,310}= 5.44$, $p < 0.05$).
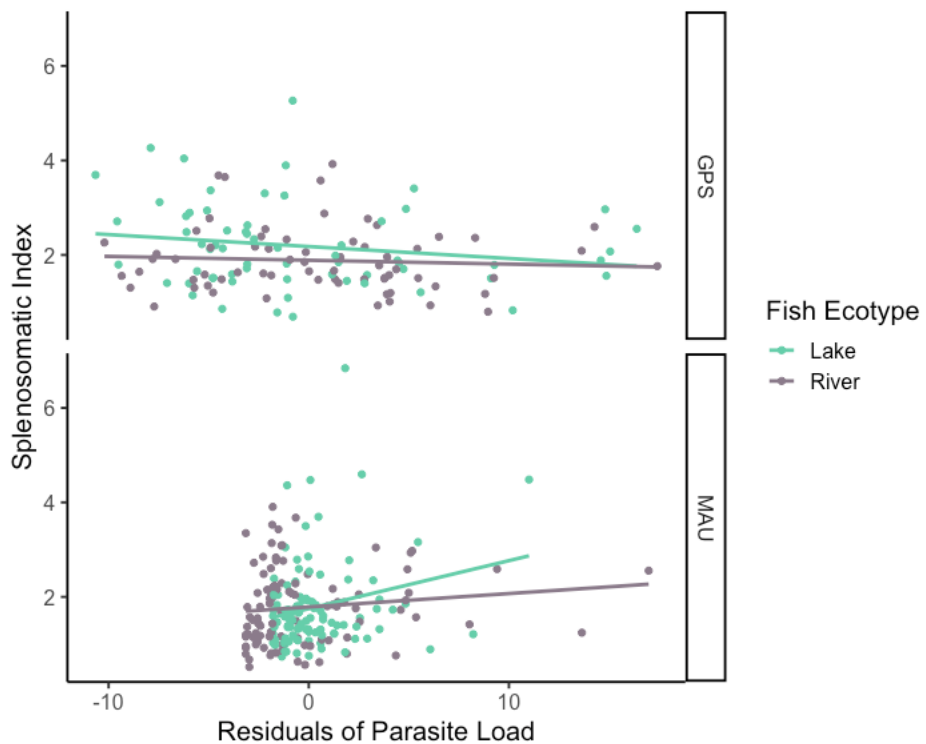
Figure 3.10: Fish splenosomatic index correlated with an interaction between parasite load, habitat of exposure and ecotype (LMM, $F_{1,322} = 10.45$, $p < 0.01$). GPS – Großer Plöner See, MAU – Malenter Au.

3.5 Discussion

Parasite-mediated selection and predation are arguably two of the strongest selection pressures driving the evolution of local adaptation (Greischar and Koskella, 2007, Kawecki and Ebert, 2004). Yet, they are not independent from each other as hosts may search for specific prey items which expose them to a specific range of parasites (Johnson *et al.*, 2009, Locke *et al.*, 2014, Stutz *et al.*, 2014). Here, we performed a field experiment focused on two parapatric river-lake stickleback population pairs translocated into a new lake and a new river habitat. We first found that Canadian fish had lower survival in the new conditions than the German fish. Furthermore, we found direct links between parasite load and long-term components of feeding ecology in the form of $\delta^{13}C$ and $\delta^{15}N$ whereby German fish were less infected and continued using optimal prey items. Noteworthy, our experimental design enabled us to estimate phenotypic plasticity in feeding ecology which revealed to be larger in lake fish than in river fish. Altogether, our study shows that feeding ecology is strongly correlated with changes in parasite load and community composition and reveal both elements of population-specificity (i.e. local adaptation) and parallel evolution.

While it is acknowledged that parasite-mediated selection and predation are important evolutionary pressures, evaluating how they simultaneously operate is difficult under laboratory conditions (Eizaguirre *et al.*, 2012b, Kaufmann *et al.*, 2015). Indeed, it is practically impossible to expose fish with the broad diversity of parasites they are exposed to in nature (Barber and Scharsack, 2010, Stewart *et al.*, 2017). And similarly, it is difficult to feed them with the broad diversity of food items they encounter under natural conditions. As such, our field experiment enables us to evaluate, in a very holistic manner, how the evolution of parasite resistance and feeding ecology operate and correlate with the evolution of local adaptation.

The first key element in the evolution of local adaptation is survival (Fraser *et al.*, 2011, Kawecki and Ebert, 2004). We found patterns of local adaptation at the country level as Canadian fish broadly had a lower survival rate than German fish. This effect was strongest under lake conditions. This survival pattern is consistent with predictions made in the local vs. foreign theory of local adaptation, whereby local genotypes outperform foreign ones (Kawecki and Ebert, 2004). This is not the first time this pattern of local adaptation is found in the three-spined stickleback (Eizaguirre *et al.*, 2012a).

The second element of local adaptation relates to individual fitness, whereby individuals will aim to maximize nutrient intake while minimizing parasite exposure and infection (Altizer *et al.*, 2018). In relation to parasite infection, results showed both elements of parallel and local adaptation. Independently of the country of origin, lake fish exposed to the Großer Plöner See lake habitat showed lower parasite load than river ecotypes from both Canada and Germany. The reciprocal result was however unclear in the river habitat. As previously speculated, patterns of parallel evolution may be clearer under lake conditions as the strength of parasite-mediated selection is stronger in that habitat (Eizaguirre *et al.*, 2012a, Eizaguirre *et al.*, 2012b, Eizaguirre *et al.*, 2011, Wegner *et al.*, 2003). Indeed, as confirmed in our experiment, lake fish are exposed to more diverse parasites than river fish (Eizaguirre *et al.*, 2012b). Different parasite communities have resulted in the evolution of specific allele pools at the genes of the major histocompatibility complex – the major immune genes of the adaptive immunity (Eizaguirre *et al.*, 2011, Wegner *et al.*, 2003). Ecotype-specific adaptation has also be detected at the level of genome-wide gene expression (Huang *et al.*, 2016, Lenz *et al.*, 2013). Noteworthy, we found evidence of local adaptation as German river fish were less infected than Canadian river fish in the lake. These results suggest there is a level of local adaptation beyond the evolution of parallel ecotypes. In the river habitat, fish are exposed to lower parasite load and diversity which may explain why patterns of parallel or local

adaptation are less clear than under lake conditions. Overall, our experiment revealed both parallel and local adaptation to parasite-mediated selection.

One of the main goals of our study was to establish the link between feeding ecology and parasite infection in the context of local adaptation. Firstly, when comparing fish from both experimental habitat types we found $\delta^{13}C$ correlated with a large four-way interaction between habitat of exposure, country of origin, ecotype and parasite load. The habitats sampled vary greatly in the prey availability and parasites taxa, and the different fish lineages have adapted to these specific differences (Brunner *et al.*, 2017, Eizaguirre *et al.*, 2011, Kalbe and Kurtz, 2006, Kalbe *et al.*, 2002, McKinnon and Rundle, 2002, Taylor and McPhail, 2000). This complex interaction however suggests genotype-by-environment interactions influence the relationship between parasites and diet. $\delta^{15}N$ was not correlated with parasite load, although we did observe a habitat of origin, country of origin, and ecotype interaction again. This suggests that variation in parasite load is not associated with the trophic level in which the stickleback feeds.

To explore the relationship further we split by habitat of exposure. Focusing on the Großer Plöner See (lake habitat) exposed fish, we found $\delta^{13}C$ correlated with an interaction between fish country of origin and parasite load, where Canadian fish with higher parasite loads fed on more benthic food sources (less negative $\delta^{13}C$) than those with lower parasite loads. A potential reason for this is that highly infected fish are trying to compensate the costs of infection by changing food source (Brunner et al 2017). A second possibility is linked to parasite-mediated behavioural responses, resulting in different feeding strategies in highly parasitised fish (Barber *et al.*, 2008, Lefevre *et al.*, 2009, Lochmiller and Deerenberg, 2000). Whilst we observed a positive correlation between $\delta^{13}C$ and parasite load in German fish, it is much weaker than that observed in Canadian fish. This may stem from adaptation at the country of origin level, where German fish are better adapted to tolerating the parasites found in the German lake

(Berner *et al*., 2008, Eizaguirre *et al.*, 2012a, Eizaguirre *et al.*, 2012b, Hendry *et al*., 2002, Kalbe *et al*., 2002, Lenz *et al*., 2013). Another potential reason behind such a finding may be that the Canadian fish are not adapted to the environmental conditions found in Germany and therefore are experiencing greater stress than the German fish, so epigenetics may play a role in the increase parasite infection and mortality (Wenzel and Piertney, 2014). It is worth noting some of the detailed patterns may also have emerged from the differential mortality of Canadian fish in the different habitats of exposure, and therefore results may need to be treated with caution.

Within Malenter Au (river habitat), we found δ^{13}C was associated with an interaction between fish ecotype and parasite load. Lake fish placed in Malenter Au showed a positive correlation between parasite load and δ^{13}C. This suggests that by feeding on their non-native prey, lake fish exposed to river conditions may acquire more parasites. Increased parasite exposure and infection may also come from the specific morphology of lake fish being less well adapted to the riverine prey species, resulting in a shift in feeding niche which could expose them to a higher diversity or abundance of parasites (Locke *et al.*, 2014, Marques *et al.*, 2011). This is supported by the lack of correlation between river ecotype parasite load and δ^{13}C. On the other hand, nitrogen isotopes were not associated with any variables tested suggesting fish continue to feed at the same trophic level regardless of parasite infection, local environment, and genetic background. It is impossible in this system to determine whether different feeding ecology results in different parasite infection, or whether different infections result in different diets. Indeed, compensatory feeding behaviours to overcome the costs of parasite infection are common in stickleback and may also happen in nature (Brunner *et al.*, 2017, Eizaguirre *et al.*, 2009b, Kalbe *et al.*, 2009). However, independent of the causal link, a suboptimal feeding niche use will correlate with increased infection and increasing selection for the evolution of local adaptation.

At this stage, it is important to note that elements of stable isotope fractionation need to be consider cautiously (DeNiro and Epstein, 1978, Fry, 2006, Minagawa and Wada, 1984, Post, 2002). Here, we tested whether fish of different origins, whether from Canada or Germany but also from lake or river origin showed different fractionation capacity. Specifically, under standardized laboratory conditions, we found that lake and river fish showed differential fraction of $\delta^{13}C$ values, even though they were fed identical food items. This possible genetic component of isotope use may explain some of the patterns detected in the field and therefore results associated with $\delta^{13}C$ and fish ecotype need to be interpreted carefully. Country of origin had no influence on fractionation of carbon. Nitrogen isotope values did not differ between any of the fish within the controlled mesocosms so we can conclude that no difference in fractionation occurs in regard to trophic level.

Thanks to the common garden nature of our experiment, we could test the level of phenotypic plasticity expressed by the fish exposed to their non-native habitat type. However, it is important to note that this study was not fully reciprocal as no mesocosms were placed in Canada, it is only reciprocal at the habitat of exposure level. We found that lake families show more change in $\delta^{13}C$ feeding adjustment values when exposed to the river conditions than the river fish exposed to the lake conditions. This indicates lake fish are capable to match levels of feeding strategy of river fish and hence have more plastic foraging capacity than their river counterparts. This effect could stem from two mechanisms. Firstly, the lake environment is more heterogeneous and sustain larger populations of fish which could result in greater genetic variation in lake ecotypes, with families capable of exploiting both the pelagic and benthic niches of the lake (Feulner *et al.*, 2015, Matthews *et al.*, 2010). Alternatively, lake fish could have higher phenotypic plasticity than river fish, independently of their niche evolution. Phenotypic plasticity has been linked to the adaptive radiation of oceanic and anadromous stickleback into

freshwater systems (Wund *et al.*, 2008, Wund *et al.*, 2012). While teasing apart the underlying mechanism is not possible in this experiment, further controlled laboratory experiments exploring feeding ability and preference as well as assimilation of food sources consumed would address this remaining knowledge gap.

Overall, our results show local adaptation of host-parasite interactions at multiple levels, country of origin, fish ecotype and habitat of exposure. The correlation between changes in parasite load and stable isotopes show that a strong relationship between parasites and feeding ecology exists, however, whether the change in diet results in the change in parasitism or whether parasitism results in shifts in diet remain to be elucidated.

Ethics statement

All animal experiments described were approved by the Ministry of Nature, Environment and Country Development, Schleswig Holstein, Germany. Permits were granted by Canadian and German governmental institutions for all steps from catching fish in Canada and bringing them to Germany.

Chapter 4: Host genetics and environment influence host-microbe interactions: A field-based common garden experiment using the three-spined stickleback (*Gasterosteus aculeatus*).

Adrienne Kerley[1], Tina Henrich[2], Christoph L. Gahr[2], Martin Kalbe[2], Lee Henry[1], China Hanson[3], Manfred Milinski[2] and Christophe Eizaguirre[1]

[1]School of Biological and Chemical Sciences, Queen Mary University of London, London, United Kingdom.

[2]Max Planck Institute for Evolutionary Biology, Department of Evolutionary Ecology, August-Thienemann-Str. 2, 24306 Ploen, Germany.

[3] Microbiology@UCL, University College London, London WC1E 6BT, UK

Keywords:

Common garden experiment, Gut microbiome, Bacteria, Local Adaptation, Host-symbiont interactions, 16S rRNA

4.1 Abstract

Identifying the determinants of a host's microbial diversity and structure is complex. Here, we explored the relative contributions of host genetics and environment by performing a field-based common garden experiment using a second-generation of lake and river three-spined stickleback (*Gasterosteus aculeatus*) from two independent systems in Canada and Germany. Fish were placed into mesocosms in either a German lake or river and were retrieved at three time points over the course of 10 months. We used 16S rRNA amplicon sequencing to characterise amplicon sequence variants (ASVs) of the fish intestinal microbiomes. To estimate the role of the environment, we quantified the fish parasite load as well as their feeding ecology using carbon (δ^{13}C) and nitrogen (δ^{15}N) stable isotopes. We found evidence that fish intestinal microbiomes are distinct from the microbiomes of their prey items and surrounding environment, showing host environmental filtering. Despite being exposed to identical environmental conditions, fish microbial diversity correlated with their country of origin, with German fish harbouring a greater diversity than Canadian fish, suggesting potential genetic determinism. Interestingly, we found evidence of host-parasite-microbe interactions, with both microbial diversity and community composition being associated with fish parasite load. Specifically, parasite infection was negatively associated with intestinal microbial diversity through interactions with fish habitat of exposure (lake or river) and country of origin (Canada or Germany). Finally, microbial diversity was correlated with fish diet (trophic position or proportion of littoral carbon) through interactions with habitat of exposure, country of origin, and fish ecotype (lake or river) showing complex host-microbe-diet interactions. Overall, our results show that it is possible to disentangle the determinant of host-microbe interactions whether linked to host genetics or environmental pressures, specifically parasite infection and diet.

4.2 Introduction

Host-microbe interactions are textbook examples of being influenced by genotype-by-environment interactions (Suzuki, 2017, Walter and Ley, 2011). However, the relative influence of these factors remains to be elucidated in natural populations (Rennison *et al.*, 2019a). Genotype-by-environment interactions are crucial pre-requisites of local adaptation, which is a common outcome of adaptive evolution, detected when resident genotypes have increased fitness in their local environment compared to non-resident genotypes (Kawecki and Ebert, 2004, Savolainen *et al.*, 2013, Sobel *et al.*, 2010).

The factors that determine host-microbe interactions can be mostly grouped into environmental and genetic categories. Environmental factors can also be further split into abiotic or biotic components. Abiotic environmental factors such as temperature or salinity correlate with host microbial diversity and community structure (Chiu *et al.*, 2020, Krause *et al.*, 2012, Lindström and Bergström, 2004, Lozupone and Knight, 2007). For instance, the microbiomes of freshwater fish are more similar to other freshwater fish than to marine ones, with *Aeromonas* and *Plesiomonas* being common in freshwater species (Nayak, 2010, Sullam *et al.*, 2012). Biotic factors, on the other hand, include elements of feeding ecology (diets) and its associated pressures such as parasite infection (Britton and Andreou, 2016). Research on the influence of diet on host microbiomes has mainly focused on humans or model organisms under controlled laboratory conditions, where hosts are fed specific dietary groups, e.g. plant-based vs. animal-based or high fat vs. low fat diets (David *et al.*, 2014, Heinritz *et al.*, 2016, Parks *et al.*, 2013). However, this is not directly representative of diet-microbe interactions in natural systems where organisms consume a wide range of food sources. Yet, replicating controlled experiments in the field is complex, especially in regard to knowing what prey item an organism has consumed. To overcome this limit, stable isotope analysis (SIA) has been used as a proxy for a long-term assessment of an individual's diet (Lorrain *et al.*, 2002, Post, 2002).

Carbon (δ^{13}C) and nitrogen (δ^{15}N) isotope ratios, in particular, can be used to calculate the proportion of littoral carbon in a host's diet, as well as their trophic position (Matthews *et al.*, 2010, Post, 2002). For instance, SIA on cichlid fish, known for their rapid local adaptation and speciation potential, showed a lack of distinct niches during the early stages of ecological diversification was associated with non-parallel changes of gut microbial communities – hence deciphering the role of diet and host genotype in the determinant of host microbiome community (Harer *et al.*, 2020).

In nature, diet is also associated with parasite infection, as different feeding strategies can expose hosts to both trophically transmitted and actively infecting parasites (Bakke *et al.*, 1992, Johnson *et al.*, 2009, Locke *et al.*, 2014, Marcogliese and Cone, 1997, Stutz *et al.*, 2014). Parasites have the ability to affect a host's microbiome (Dheilly *et al.*, 2015, Fredensborg *et al.*, 2020, Llewellyn *et al.*, 2017, Rausch *et al.*, 2013), as studies have shown both increased (Lee *et al.*, 2014, Rosa *et al.*, 2018) and decreased microbial diversity in association with parasite exposure (Houlden *et al.*, 2015, McKenney *et al.*, 2015). This may be a response to the influx of parasite-associated microbes, as each parasite harbours their own unique microbiome (Dheilly *et al.*, 2015). Parasites also have the ability to directly modify the microbiome and alter host immune response to create more favourable conditions for their own growth (Dheilly *et al.*, 2015). Interestingly, the microbes within a host are not passive in this relationship, with the ability to both facilitate infection and assist the host in pathogen defence (Ford and King, 2016, King and Bonsall, 2017, King *et al.*, 2016, Stevens *et al.*, 2021). It is widely accepted that local adaptation of the host can emerge as a response to local parasite communities (Eizaguirre *et al.*, 2012a, Hamley *et al.*, 2017, Kalbe and Kurtz, 2006, Kaufmann *et al.*, 2017, Lenz *et al.*, 2013, Summers *et al.*, 2003, Weber *et al.*, 2017), and growing research supports the theory that host-parasite-microbe interactions could also be associated with local adaptation of a host population (Dheilly *et al.*, 2015, King *et al.*, 2016, Kwiatkowski *et*

al., 2012). To properly test this hypothesis, however, it is important to understand the genetic determinants of host's selection on microbial communities (Smith *et al.*, 2015, Steury *et al.*, 2019, Sullam *et al.*, 2012, Wang *et al.*, 2016a).

Indeed, microbiomes of genetically similar hosts, for instance at the genes of the major histocompatibility complex (MHC), are more similar than those of unrelated hosts exposed to similar environments (Bates *et al.*, 2007, Bonder *et al.*, 2016, Chen *et al.*, 2018, Steury *et al.*, 2019, Zoetendal *et al.*, 2001). Host genetics can also influence their microbiome in other ways such as behavioural differences where they will avoid areas deemed unsanitary or containing rancid food sources, which will limit immigration of potentially harmful bacteria (Welzl *et al.*, 2001). A host can also target the growth of bacteria that is beneficial to them, such as those that provide nutrients the host is unable to get through their diet by providing nutrients and suitable living conditions for those bacteria species (Sonnenburg *et al.*, 2005). A fascinating form of genetic control is when a host is able to monitor the location bacteria or the benefits provided by those bacteria within their microbiome through their toll-like receptors, implementing further controls if required (Kiers *et al*., 2003, Vaishnava *et al*., 2011). For example, legumes harbour *Bradyrhizobium japonicum* which produce nitrogen for the host, these are compartmentalised in the plants root nodules and given nutrients to survive. However, if the host senses that the bacteria stop producing nitrogen they will cut the nutrient supply to that root nodule (Kiers *et al*., 2003). Similarly, as individuals of different sexes exhibit different behaviours or may use/process food sources differently, sex genotypes have been associated with sex-specific microbiomes (Bolnick *et al.*, 2014c, Dominianni *et al.*, 2015, Zhang *et al.*, 2020).

Resolving whether host environment or genetics underpins microbiome structure and association with host local adaptation is challenging. To overcome this problem, laboratory experiments have been successfully used (Douglas, 2019, Spor *et al.*, 2011).

However, evidence suggesting captivity can affect the host microbiome, coupled with the lack of realistic complexity of laboratory experiments, shows sampling wild populations is crucial to understanding host-microbe interactions fully (Hird, 2017, McKenzie *et al.*, 2017). A replicated common garden experiment is an ideal way to study the influence of genotype-by-environment interactions and local adaptation on host microbiomes in the wild (de Villemereuil *et al.*, 2016). This allows the study of 'local vs. foreign' criterion, i.e. how hosts fitness from different lineages vary within the same habitat, here with a special focus on observing the influence of genetic factors on the microbiome. We can also study the 'home vs. away' criterion, i.e. how hosts from the same linage vary across two different habitats, to explore the role the local environment has on host microbiomes (Kawecki and Ebert, 2004). Additionally, signals of local adaptation, such as survival rate, host growth, and body condition, can be measured to observe whether fitness varies between local and foreign genotypes. Overall, to test for the role of microbiome in host local adaptation, it is important to focus on species for which local adaptation has been well described.

The three-spined stickleback (*Gasterosteus aculeatus*) is an ideal model organism to explore whether locally adapted fish populations have population-specific microbiomes, i.e. stems from genetic determinism, or whether the microbial communities are mostly determined by local environmental pressures. Since the last glaciation, multiple colonisations of freshwater habitats have resulted in natural parallel systems of lake-river ecotypes (Bell and Foster, 1994, Eizaguirre *et al.*, 2011, Feulner *et al.*, 2015, Marques *et al.*, 2016). These ecotypes differ in genetics, morphology, physiology and behaviours (Bolnick *et al.*, 2018, Cano *et al.*, 2006, Eizaguirre *et al.*, 2011, Hanson *et al.*, 2017, Marques *et al.*, 2016, Ravinet *et al.*, 2013, Rennison *et al.*, 2019b, Taylor and McPhail, 2000). They also experience variation in parasite infections, for example lake fish are exposed to a greater parasite diversity and load than river fish (Bolnick *et al.*,

2020, Eizaguirre *et al.*, 2011, Kalbe *et al.*, 2002, Reusch *et al.*, 2001). As a consequence, lake ecotype populations have evolved better resistance to more diverse, higher parasite loads than river populations (Eizaguirre *et al.*, 2012a, Eizaguirre *et al.*, 2012b, Kalbe and Kurtz, 2006, Lenz *et al.*, 2013, Wegner *et al.*, 2003). Additionally, lake-river ecotype pairs have different feeding strategies (Berner *et al.*, 2008), which likely alter exposure to parasites (Stutz *et al.*, 2014) and influence the microbial community (Bolnick *et al.*, 2014b).

Studies on the determinants of stickleback gut microbiome identified that fish genetics, diet, helminth infection, sex, ecotype and polymorphism in the MHC gene correlated with microbial diversity and community structure (Bolnick *et al.*, 2014a, Bolnick *et al.*, 2014b, Bolnick *et al.*, 2014c, Ling *et al.*, 2020, Smith *et al.*, 2015, Steury *et al.*, 2019). Additionally, parallel evolution of the gut microbiome has been observed across independently evolved benthic and limnetic stickleback ecotype pairs from Canada (Rennison *et al.*, 2019a). However, no studies have used replicated common garden experiments on the three-spined stickleback to test for local adaptation of host microbiomes or to explore genotype-by-environment interactions.

Here, we conducted a reciprocal common garden experiment by transplanting both Canadian and German lake and river three-spined stickleback into a third lake and river system in Germany. We tested whether fish origin, the habitat of exposure and parasite infection correlated with the diversity and composition of stickleback gut microbiomes. Additionally, we tested for signals of local adaptation whereby the presence of certain bacterial species would be associated with increased fish fitness proxies. Finally, we tested whether diet correlated with microbial diversity. For this, we used the proportion of littoral carbon used and fish trophic position calculated from stable isotope ratios as direct estimators of diet. This *in situ* common garden experiment allowed us to explore the influence of host environment and genotypes on the stickleback gut

microbiome without the constraints associated with laboratory experiments. This study was conducted in a broad framework where water and fish prey items were also collected and sequenced to determine their microbial community diversity.

4.3 Methods

4.3.1 Breeding design

Three-spined sticklebacks were collected using hand nets and minnow traps from two paired lake and stream populations, one in Canada (McCreight lake: 50˚28'12.4''N, 125˚65'31.7''W, Amour de Cosmos creek: 50˚23'54.3''N, 125˚63'62.9''W) and one in Germany (Westensee lake: 54˚26'89.8''N, 9˚96'09.2''E, Eider stream: 54˚16'65.5''N, 10˚07'60.1''E). Fish were collected during two breeding seasons, 2014 and 2015. A total of 20 individuals from each sample population were used to breed first-generation (G1) families through *in vitro* fertilisation. To stop fungal growth, eggs were treated post-fertilisation with acriflavine (Dajana) and Methylene Blue (King British). Fertilised Canadian eggs were held at 4˚C during transportation to Germany. Transportation conditions were replicated for German fish by storing fertilised eggs in the fridge at 4˚C for 4 days. G1 eggs were then incubated, hatched and fish were raised in tanks in the laboratory with constant water flow, fed frozen chironomid larvae *ad libitum* and held in controlled summer conditions at 18˚C, 18:6 Light:Dark (L:D). Sexual maturity was triggered in fish by cycling artificially through the seasons (Autumn: 12˚C for 2 weeks, 12:12 L:D, Winter: 6˚C for 4 weeks, 12:12 L:D, Spring: 12˚C for 4 weeks, 12:12 L:D and Summer: 18˚C until breeding, 18:6 L:D). Nest materials, green polyester threads (cut to a length of ~ 10 cm) and a sand-filled petri dish, were provided to sexually mature males kept in individual tanks. To initiate natural spawning, an unrelated gravid female from the same sampling population was placed into the tank with the male. Breeding pairs were allowed to mate repeatedly to create family groups of identical genetic origin. We obtained a total of 24 second-generation (G2) families (N = 12 of Canadian origin and N

= 12 of German origin). Using a G2 fish generation enabled us to reduce the influence of possible parental effects. All experiments described were approved by the Ministry of Nature, Environment and Country Development, Schleswig Holstein, Germany. Permits were granted by Canadian and German governmental institutions for all steps of wild fish collection.

4.3.2 Field-based common garden experiment

Before placing the fish into experimental mesocosms in a lake and a river, dorsal spine clippings were collected from each individual. DNA extraction on the spine clipping was conducted using DNA Tissue kit (Invitek, Germany) following manufacturer's protocols to allow for sex-typing and genotyping for 14 microsatellites for later identification of individuals (Kalbe *et al.*, 2009). At that time, fish standard length and weight were also recorded to measure individual growth over the course of the experimental period. In October 2016, 48 mesocosms made of 5 mm stainless steel mesh and a stable framework (L:1 m, H:0.25 m, W:0.6 m) were placed in two geographically connected exposure habitats, a large lake Großer Plöner See (GPS, $N_{mesocosms}$ = 24, 54˚14'61.0''N, 10˚40'86.9''E) and Malenter Au a small, slow-flowing stream (MAU, $N_{mesocosms}$ = 24, 54˚19'62.7''N, 10˚55'65.9''E). Field exposure protocols match those described in Eizaguirre *et al.* (2012a). We chose this German lake-river system (Großer Plöner See - Malenter Au) as the location for our experimental mesocosms as it is isolated from the lake-river system (Westensee lake - Eider stream) that parental German fish were collected from. This assured that both Canadian and German fish were exposed to a different lake-river system than their natal one, guaranteeing all fish were in an allopatric system, enabling more direct comparisons between fish of different origins to be made. Mesh size allowed for constant water flow and invertebrates (food sources and intermediate hosts for parasites) to pass through, whilst keeping experimental fish separate from wild stickleback and predators. Malenter Au mesocosms were placed 3 m

apart at a depth of 0.5 – 1.5 m. Großer Plöner See mesocosms were placed 2 m apart at 1 – 1.2 m depth. Each mesocosm contained 24 fish of either Canadian or German origin, of which 12 were river ecotype fish (♀= 6, ♂= 6) and 12 were lake ecotype fish (♀= 6, ♂= 6, Figure 4.1). Mating was prevented by separating sexes within the mesocosm with a plexiglass. Fish were kept in mesocosms for up to 10 months and were sampled at three different time points during this experimental period. The first collection time point occurred in December 2016, the second in May 2017 and the final time point in July 2017. Predicting increasing mortality rates throughout the experiment, we dedicated 25% of the mesocosms to the December time point, another 25% to the May time point and the remaining 50% of mesocosms were sampled during the final (July) time point. Experimental sites were visited weekly to check for dead fish and guarantee that mesocosms remained submerged. Six additional mesocosms were kept in controlled biotic and abiotic conditions within the laboratory habitat, referred to as the control mesocosms. These mesocosms were cycled through the seasons to mimic external environmental conditions with controlled temperature and light periods (specific details for each season given above). Laboratory fish were fed *ad libitum* on frozen chironomid larvae.
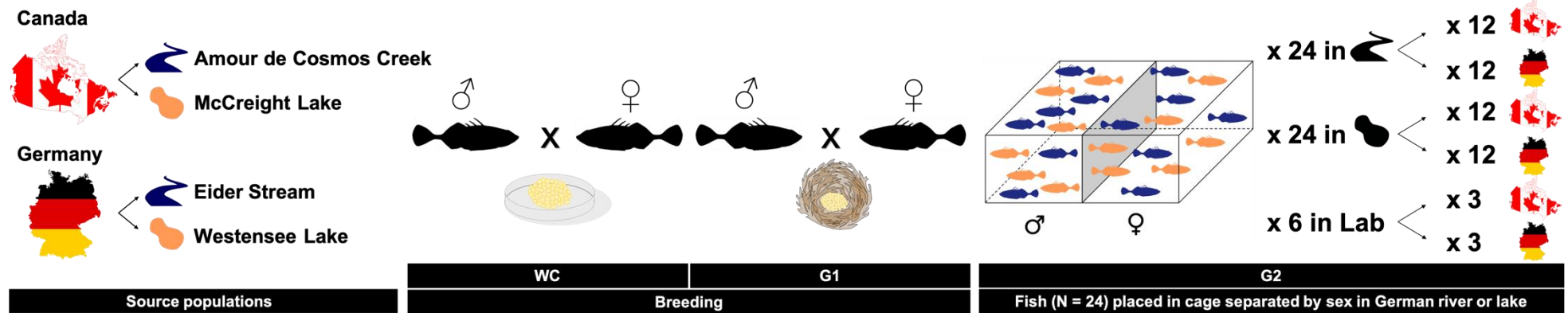
Figure 4.1: Experimental design of the field-based common garden experiment. Wild-caught (WC) fish were collected from rivers and lakes (coloured to depict fish ecotype, blue = river ecotype, orange = lake ecotype) in Canada (Amour de Cosmos creek and McCreight lake) and Germany (Eider stream and Westensee lake) were bred in vitro. First-generation (G1) fish were allowed to breed naturally to produced second-generation fish (G2). G2 fish were placed in mesocosms in a new geographically connected river and lake habitat in Germany (Malenter Au and Großer Plöner See). Mesocosms were stocked with either Canadian or German fish and contained 24 fish (6 female river fish, 6 female lake fish, 6 male river fish and 6 male lake fish). Sexes within a mesocosm were separated with plexiglass to prevent mating. 25% of mesocosms were collected in December 2017, 25% in May 2017 and all remaining fish were collected in July 2017.

4.3.3 Sample collection

Twelve mesocosms from both habitats of exposure were retrieved in December and May, and 24 mesocosms July, as well as the six control mesocosms from the laboratory. Dissection occurred within two days of sampling. Fish were euthanised with tricaine methanesulfonate (MS222, 200mg/l, Sigma). External surfaces of fish were screened for ectoparasites under a dissection microscope. Eyes, gills and internal organs, except the spleen and the heart, were screened for endoparasites, using a sterilised compressorium (Kalbe *et al.*, 2002). Parasite abundance and diversity were combined into an individual parasite index, I_{PI} (Kalbe *et al.*, 2002). I_{PI} values were calculated for all fish across both habitats of exposure, as well as separating for Großer Plöner See and Malenter Au exposed fish. Standard length (SL), total length (TL), total weight, and organ weight (spleen and head kidney) were measured. Proxies of fish fitness were estimated using body condition ($CF = \left(\frac{Body\ Weight}{Standard\ Length}\right)^3 \times 100$) (Frischknecht, 1993) and splenosomatic index ($SSI = \frac{Spleen\ Weight}{Body\ Weight} \times 100$) (Kalbe *et al.*, 2009). Intestines were stored in a physiological saline solution (0.64% sodium chloride) at -80°C until DNA extraction.

During the July time point, invertebrate samples were collected for microbial analysis to compare fish gut microbiomes to that of their food source. A plankton net was used to sample invertebrates in Großer Plöner See and Malenter Au. Invertebrates were transferred to a sterile falcon tube to be sorted in the laboratory. All invertebrates observed that were potential food sources of the three-spined sticklebacks were selected with sterile forceps (5 individuals per invertebrate species) and stored in ethanol-filled, sterile microcentrifuge tubes at −80 °C until DNA extraction. Invertebrates used to feed the laboratory fish were aseptically removed from their packaging, defrosted and placed into individual sterile microcentrifuge tubes containing ethanol using sterile forceps. Invertebrate samples obtained for microbiome samples included caseless caddisfly larvae, mayfly larvae, isopods, *Gammarus* sp. and white chironomid larvae from

Malenter Au, as well as *Gammarus* sp. and red chironomid larvae from Großer Plöner See. White and red chironomid larvae were obtained from the laboratory.

Water samples were also collected in July. Five replicates (5 x 500 ml) were collected for each habitat of exposure, in sterile Duran® bottles, and transported to the laboratory on ice. Water samples were filtered using a Thermo Scientific Nalgene Polysulfone Filter Holder and Receiver (500 ml capacity). All filtration equipment was sterilised by autoclaving immediately before filtration. Water samples were first passed through a sterile 0.7 μm Whatman® glass fibre filter to remove large organic matter. Pre-filtered water was then passed through a sterile 0.22 μm Whatman® nitrocellulose membrane filter. Filters were transferred to sterile Falcon tubes and stored at −80 °C until DNA extraction. A filter blank was taken at the start and end of processing the replicates of each habitat of exposure by loading a 0.22 μm Whatman® nitrocellulose membrane filter and passing through 500 ml of sterile distilled water (Milli-Q® water). Aseptic techniques were used throughout the filtration process.

4.3.4 DNA extraction

DNA extractions were carried out in a class II biological safety cabinet to reduce the risk of contamination. Fish intestines were cut in half to reduce the risk of overloading the spin column. To remove excess salt (an artefact of the chosen storage method) from fish samples prior to DNA extraction, cut intestines were placed in sterile microcentrifuge tubes containing 850 μl of 100% ethanol at -20°C overnight, followed by centrifuging for 12 minutes at 6000g, the supernatant was then discarded. A second rinse in 850 μl of 70% ethanol was carried out, then samples were centrifuged for 12 minutes at 6000g and the supernatant was discarded. Intestines were left for 10 minutes in the class II biological safety cabinet to make sure ethanol had evaporated prior to starting the DNA extraction process. DNA was extracted from samples using DNeasy 96 Blood & Tissue Kit (QIAGEN) with the following protocol modifications. For invertebrate samples, whole

individuals were used. For water samples, nitrocellulose membrane filters were torn into small pieces using sterile forceps to allow for the whole sample to be submerged in the lysis buffer. For all samples, lysis was carried out overnight and the elution step was conducted twice with 75 μl of warmed elution buffer (50°C) incubated for 5 minutes each time before centrifuging, to increase DNA yield. Each 96 Blood & Tissue plate contained two blanks to check for contamination at each stage of processing, one to be used as an extraction blank and the second as a PCR blank.

4.3.5 PCR amplification and sequencing

The V4 16S ribosomal RNA gene (390 bp) was amplified using the primers 515F (5'-GTGYCAGCMGCCGCGGTAA-3') – 806R (5'-GGACTACNVGGGTWTCTAAT-3', (Apprill *et al.*, 2015, Parada *et al.*, 2016). PCR reactions were carried out in duplicate to reduce the risk of PCR artefacts or failure. Each reaction consisted of 5 μl 5x PCRBIO HiFi buffer, 1 μl of forward and reverse primers (5 pmol/μl), 0.25 μl PCRBIO HiFi Polymerase (2 u/μl), 16.75 μl PCR grade H_2O and 1 μl Template DNA for a total volume of 25 μl. Reactions were conducted on a Bio-Rad T100 Thermal Cycler using the following protocol: 1 minute at 95°C, 15 seconds at 95°C (x34), 15 seconds at 65°C (x34), 30 seconds at 72°C (x34), 5 minutes at 72°C. Amplification was validated using gel electrophoresis. PCR duplicates were pooled prior to sequencing. Extraction blanks were treated the same as samples during amplification. The only variation in the protocol for PCR blanks was the substitution of 1 μl Template DNA with 1 μl PCR grade H_2O.

4.3.6 Illumina sequencing data analysis

The Genome Centre, London, carried out PCR product clean up, sequence library prep and Illumina MiSeq 300bp paired-end sequencing. FastQC was used to observe the quality of demultiplexed paired-end sequencing reads (R1 and R2, Andrews, 2010). Primer and adapter sequences were removed using Trimmomatic (Bolger *et al.*, 2014). We estimated trimming parameters using Figaro v1.0.0, a bioinformatics tool that allows

for post-trimming sequence information to be maximized, whilst expected errors in the sequences themselves are minimized (White *et al.*, 2008). An amplicon size of 390bp and a minimum overlap of 70bp, to allow for the trimming of up to 100bp from the overlapping region, were used as Figaro parameters. This resulted in 28bp and 75bp being removed from the 3' end of the sequence for R1s and R2s respectively. QIIME 2 v2020.02 (Bolyen *et al.*, 2019) and DADA2 software packages (Callahan *et al.*, 2016) were used to denoise, merge R1s and R2s, remove chimeras, and assign amplicon sequence variants (ASVs) using default parameters. Taxonomy was assigned using Greengenes 13_8 database (McDonald *et al.*, 2012). Previous microbiome studies used operational taxonomic units (OTUs) to cluster bacterial sequences by sequence similarity (typically at a 97% similarity threshold), however the field has now shifted to using ASVs as the preferred method as they are exact sequence variants that are comparable across studies, improving reproducibility (Callahan *et al.*, 2017).

4.3.7 Stable isotope analysis

For a subset of fish (N = 268) from the July time point, we sampled a small section of white muscle from the tail and stored them at -20°C for stable isotope analysis. Additionally, mussels (N = 3) and snails (N = 3) were collected from Großer Plöner See and Malenter Au during invertebrate sampling and frozen at −20 °C for stable isotope analysis. These filter feeders and grazers were used as baselines for stable isotope analysis and allowed for fish trophic position and proportion of littoral carbon to be calculated. Stable isotope samples were desiccated at 60°C for 48 hours and homogenised using a mechanical grinder for 1 minute (30Hz, 2 x 316 stainless steel beads, TissueLyser II, Qiagen). Mussel and snail samples were removed from their shells before drying. Approximately 1 mg of homogenised sample was loaded into tin capsules (6 x 4 mm). We measured carbon and nitrogen isotopes. Samples were analysed on an Integra 2 spectrometer (Sercon Instruments, Crewer, UK, Analytical precision: 0.1‰) using

continuous-flow isotope ratio mass spectrometry (CF-IRMS). Isotope ratios were calculated for carbon (δ^{13}C) and nitrogen (δ^{15}N). Isotopic ratio accuracy was verified using a Protein (Casein) Standard OAS/Isotope every 10 samples (Elemental Microanalysis, δ^{13}C: -26.98 ± 0.13‰, δ^{15}N: +5.94 ± 0.08‰).

Proportion of littoral carbon (α) within an individual's diet was calculated by comparing fish δ^{13}C with the average δ^{13}C of the primary consumers within the system $\left(\alpha = \frac{\delta^{13}C_{fish} - \delta^{13}C_{mussels}}{\delta^{13}C_{snails} - \delta^{13}C_{mussels}} \right)$. In freshwater systems, filter feeders such as mussels and clams represent pelagic baselines, whilst grazers such as snails are the littoral baselines (Post, 2002). Nitrogen isotope ratios are enriched by ~3.4 ‰ in consumers in comparison to their food source, allowing for individual trophic position (tpos) to be calculated (Post, 2002). To calculate tpos, the baseline δ^{15}N of a system is required ($\delta^{15}N_{baseline} = \alpha\delta^{15}N_{snails} + (1 - \alpha)\delta^{15}N_{mussels}$), this value can then be used to calculate an individual's trophic position $\left(tpos = 2 + \frac{1}{3.4}\left(\delta^{15}N_{fish} - \delta^{15}N_{baseline}\right) \right)$.

4.3.8 Statistical analyses

Statistical analyses were conducted in R version 3.6.3, packages used include: *phyloseq, vegan, lmerTest, lme4* (McMurdie and Holmes, 2013, R Core Team, 2020). The normality of model residuals were tested and data transformation was conducted, if required, to meet test assumptions. Collinearity between fixed factors within a model was tested using linear models (LM) and if observed, residuals of their regressions were used. All models were backward selected using Akaike's Information Criterion (AIC) values to retain the optimal reduced model (Burnham and Anderson, 2004). Restricted Maximum Likelihood was used for all linear mixed effect models due to the slight variation in sample numbers due to fish mortality during the course of the experiment. To test for correlations between parasite load (I_{PI}) and microbial community composition, an infection category was assigned to each fish in relation to their I_{PI}. Fish with an I_{PI} in the

lower tertile were classified as having a low infection, those in the upper tertile were classified as having high infections, the remaining individuals were classified as medium.

4.3.8.1 Amplicon sequence variants

ASVs classified as mitochondria or chloroplast, as well as ASVs with fewer than ten reads across at least two samples, were removed from the dataset. Additionally, samples with fewer than 500 reads were removed from our dataset. Analyses were run on non-rarefied data and data were rarefied to 1000 reads to test for the consistency of the detected patterns.

Understanding whether similar bacteria are common across all sample types and habitats of exposure can help to highlight the influence of the environment on microbes. Bacteria, at the ASV and phylum level, were classified as shared, partially shared or unique between sample types, including fish gut, invertebrate and water samples. Fish-specific shared, partially shared or unique ASVs and phyla were also identified across the three habitats of exposure (Großer Plöner See – lake, Malenter Au – river, and laboratory). Additionally, we identified core ASVs within fish gut samples across the three exposure habitats. ASVs were classified as core microbiome if they were present in at least 65% of fish within a habitat of exposure.

4.3.8.2 Testing for survival differential

Whether an organism survives in an environment is essential for the evolution of local adaptation. To test fish survival, we used a binomial generalised linear mixed effect model (GLMM) to test whether survival was associated with their country of origin (Canada or Germany), habitat of exposure (Großer Plöner See or Malenter Au) or ecotype (lake or river) and the interaction of these three variables, with mesocosm ID, collection month and sex set as random effects. Control mesocosms were tested separately as they

were sampled at a single time point. Tukey post-hoc tests were carried out to examine pairwise comparisons between categories of the significant variables.

4.3.8.3 Do different sample types have different microbiomes?

To determine the role of environment in influencing the diversity of microbes in fish guts, it is important to know how they compared to the microbial diversity found in invertebrates and water. To explore this question, we fitted linear mixed effects models (LMM) to test for associations between Shannon Diversity, Faith's Phylogenetic Diversity (PD) and Gini-Simpson indexes, with sample type and habitat of exposure, as well as their interaction. Due to invertebrate and water samples only being collected in July, we used collection month as a random effect. A Tukey post-hoc test between sample type and the habitats of exposure examined pairwise comparisons between significant factors.

To investigate whether microbial community composition differed among the sample types and their habitat of exposure, we used PERMANOVAs based on unweighted UniFrac distance, weighted UniFrac distance and Bray-Curtis dissimilarity. Collection month was set as a block using strata (i.e. random factor for a PERMANOVA). Pairwise PERMANOVAs were calculated between significant factors using the *pairwise.adonis2* function from the pairwiseAdonis package. All significant effects were followed by a dispersion test with *betadisper*, from the vegan package, to ascertain the observed differences were not related to heterogeneity in dispersion.

4.3.8.4 Do fish held in different habitats harbour different microbiomes?

After establishing the difference between the field-based fish, laboratory fish and different sample types, we focused on what influences stickleback gut microbiome diversity and composition under wild conditions. Firstly, the determinants of variation in fish microbial diversity linked to habitat of exposure, country of origin, ecotype, family

background but also month of sampling and mesocosm ID were identified using a variance component analysis from the VCA package (Schuetzenmeister and Dufey, 2019).

Then, linear mixed effects models (LMM) were used to test for differences in microbial diversity (Shannon, PD, Gini-Simpson) associated with fish habitat of exposure, country of origin, ecotype, I_{PI} and sex, as well as their two and three-way interactions. Mesocosm ID and month of collection were used as random effects. Tukey post-hoc tests between significant factors were used for pairwise comparisons.

PERMANOVAs were used to test whether fish habitat of exposure, country of origin, ecotype and I_{PI} group, as well as their two and three-way interactions correlated with microbial community composition across the three β-diversity metrics. Collection month was used as a block. Pairwise PERMANOVAs were calculated between significant factors. All significant effects were followed by a dispersion test. A similarity percentage analysis (SIMPER) was conducted to identify ASVs that significantly contribute most to variation among significant groups using the *simper.pretty* function (Steinberger, 2018), followed by Kruskal-Wallis tests with false discovery rate (FDR) corrected p-values using the function *kruskal.pretty* (Steinberger, 2018).

4.3.8.5 Do fish within a singular habitat have different microbiomes?

Due to the strong effect that the habitat of exposure has on microbial diversity and composition, we split the dataset for the two habitats of exposure, Malenter Au and Großer Plöner See. To determine differences in microbial diversity within the two separate habitats, linear mixed effects models (LMM) that included fish country of origin, ecotype, I_{PI} and sex as well as their two and three-way interactions as fixed predictors were used. We assigned mesocosm ID and month of collection as random effects. Tukey post-hoc tests were carried out on significant factors.

PERMANOVAs were used to test whether fish country of origin, ecotype and I_{PI} group as well as their two and three-way interactions correlated with microbial community composition (all β-diversity metrics). Month of collection was used as a block. Pairwise PERMANOVAs were calculated for significant effects. All significant effects were followed by a dispersion test. A SIMPER was conducted to identify ASVs that significantly contribute most to variation among significant groups followed by Kruskal-Wallis tests.

4.3.8.6 The influence of feeding ecology on the gut microbiome

Diet is known to have a strong influence on the gut microbiome of a host (David *et al.*, 2014). Therefore, we tested for correlations between microbial diversity (Shannon, PD and Gini-Simpson), fish diet, i.e. littoral carbon and trophic position, habitat of exposure, country of origin, ecotype, sex as well as their two and three-way interactions using LMMs. Mesocosm ID was used as a random effect. Stable isotope samples were only collected in July so month of collection was not required as a random effect. Proportional use of littoral carbon was expressed as the residuals of the regression between littoral carbon, habitat of exposure and fish trophic position. Residuals were also used for fish trophic position as this variable was collinear with habitat of exposure. We did not test the habitats of exposure separately for this part of the analysis as we were interested in the interaction between habitat of exposure and other factors tested.

Next, we explored the influence of fish diet, littoral carbon and trophic position, on the relative abundance of ASVs. Only common ASVs that had a mean relative abundance of $> 0.01\%$ were selected. We used quasibinomial generalised linear mixed models (GLMS) to evaluate the association between individual ASV relative abundance and both fish proportional use of littoral carbon and trophic position. Chi-squared tests were used to identify whether the number of significant models obtained exceeded the

5% null expectation. We applied FDR analysis to account for false positives (type I error) arising from multiple statistical comparisons to obtain the number of significant models.

4.3.8.7 Microbes and fish fitness

Finally, we tested whether fitness traits were correlated with an interaction between individual ASVs and habitat of exposure. We used a similarity percentage analysis (SIMPER), followed by Kruskal-Wallis tests, to identify ASVs of interest that significantly contributed most to variation among habitat of exposure between the country of origin groups. Fish were split into their country of origin as previous tests showed microbiomes of fish from Canada and Germany differed within the same habitat. Linear models were used to identify whether the interaction between ASVs of interest and habitat of exposure correlated with fish fitness traits. Fish traits tested were standard length at the end of the experiment (SL), body condition (CF) and splenosomatic index (SSI) as a proxy for immune activation.

4.4 Results

4.4.1 Fish survival differential

Out of 1296 fish used at the start of the experiment, from 54 experimental mesocosms in Großer Plöner See and Malenter Au, as well as in the control mesocosms, 1000 fish survived (Supplementary Table 4.5). Within the experimental habitats, we found the likelihood of survival correlated with an interaction between country of origin and habitat of exposure (GLMM, $X^2_2 = 8.02$, p < 0.01). Canadian fish placed within Großer Plöner See had a reduced likelihood of survival in comparison to Canadian fish within Malenter Au or German fish in either habitat of exposure (TukeyHSD, p < 0.001, Figure 4.2, Supplementary Table 4.6). Fish survival was not correlated with ecotype (GLMM, $X^2_1 = 0.93$, p = 0.35). Fish held in controlled laboratory conditions also showed that country of origin correlated with the likelihood of survival (GLMM, $X^2_1 = 5.14$, p <

0.05). However, in the laboratory German fish had a decreased likelihood of survival in comparison to Canadian fish (Figure 4.2).
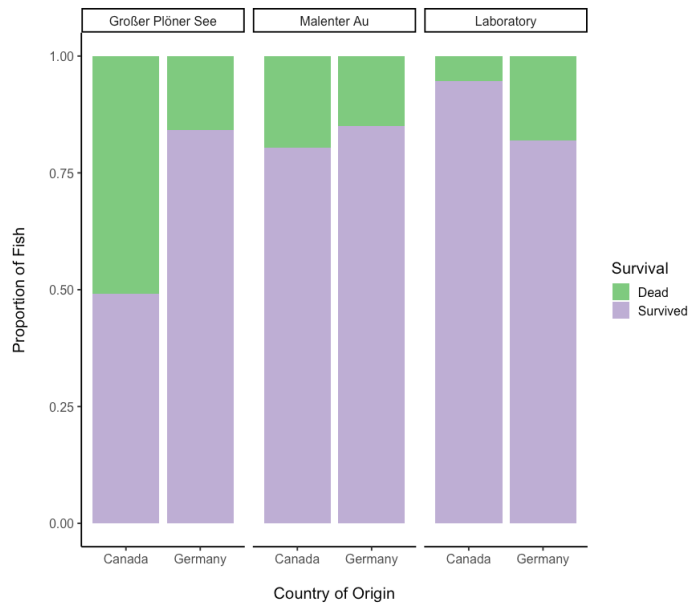
Figure 4.2: Fish likelihood of survival differed with an interaction between habitat of exposure and country of origin (GLMM, Experimental mesocosms, GPS and MAU: $X^2{}_1 = 8.02$, $p < 0.01$, control mesocosms: $X^2{}_1 = 5.14$, $p < 0.05$).

4.4.2 Microbiome descriptive summary

After removing primers and chimeras, 3,511,703 reads were retained from the 751 sample sequences. A total of 309 samples were removed as they had fewer than 500 reads, retaining 3,211,982 reads across 442 samples, of which 120 also had stable isotope data. Additional analysis was carried out on samples rarefied to 1000 reads (N = 364), and both non-rarefied and rarefied datasets were used for core microbiome identification, bacterial diversity and community composition analysis. Because of overall similarity in patterns, only non-rarefied results are reported (but see Supplementary Materials – Chapter 4 for rarefied results). Rarefaction curves for the non-rarefied and rarefied datasets showed they reached saturation plateaus (Supplementary Figure 4.1). Extraction and PCR blanks contained extremely low read numbers, showing the reliability of the protocols.

A total of 30 different phyla were identified across all sample types, with fish showing a large amount of inter-individual variation (Supplementary Figure 4.2). Within

fish samples, Proteobacteria was one of the most dominant phyla with a mean relative abundance of 58 ± 1% (standard error, SE). Firmicutes were the next most dominant phyla (13 ± 1%), followed by Planctomycetes (10 ± 1%). Proteobacteria and Firmicutes were also dominant within invertebrate samples (63 ± 3% and 5 ± 1%), as well as Bacteroidetes (18 ± 2%). Similarly, water samples also had a high relative abundance of Proteobacteria (31 ± 5%) and Bacteroidetes (32 ± 10%), as well as a high abundance of Actinobacteria (34 ± 6%).

4.4.3 Amplicon sequence variants

We identified 2562 ASVs across all samples. A total of 1639 ASVs were ubiquitous across all sample types. Fish samples contained 62 fish-specific ASVs, whilst invertebrates had 59 and 75 ASVs were found only in water samples (Figure 4.3A). When focusing on bacterial phyla, we found 28 of the 30 phyla sequenced were present in all sample types (Figure 4.3B). The phylum OP1 was only found in water samples. A second phylum, NC10, was unique to invertebrate samples. No bacterial phyla were unique to fish samples.

We found a similar mix of shared and unique ASVs when comparing fish from different habitats of exposure. Fish contained 2201 ASVs, 305 of which were ubiquitous across all habitats of exposure (Figure 4.3C). Fish held within Malenter Au harboured 737 unique ASVs in comparison to Großer Plöner See's 222 unique ASVs and laboratory fish's 207 ASVs. At the phylum level, fish from Malenter Au contained two unique phyla, Synergistetes and SR1, whilst Großer Plöner See and laboratory fish contained no unique phyla (Figure 4.3D).

When we examined the core gut microbiome of fish across the three habitats of exposure, two ASVs, a *Ralstonia* species and an ASV belonging to the family Oxalobacteraceae, were identified as core in all exposure habitats at a 65% prevalence threshold. *Propionibacterium acnes* and a *Rickettsiella* species (*Rickettsiella* species 1)

were present in at least 65% of Malenter Au fish. Fish within Großer Plöner See had three core ASVs, a *Synechococcus* species, a *Rickettsiella* species (*Rickettsiella* species 2) and an ASV belonging to the Isosphaeraceae family. An additional five core ASVs were identified in control fish from the laboratory, two ASVs belonging to the Rhizobiales order, two ASVs belonging to the Bacillaceae family and *Reyranella massiliensis*.
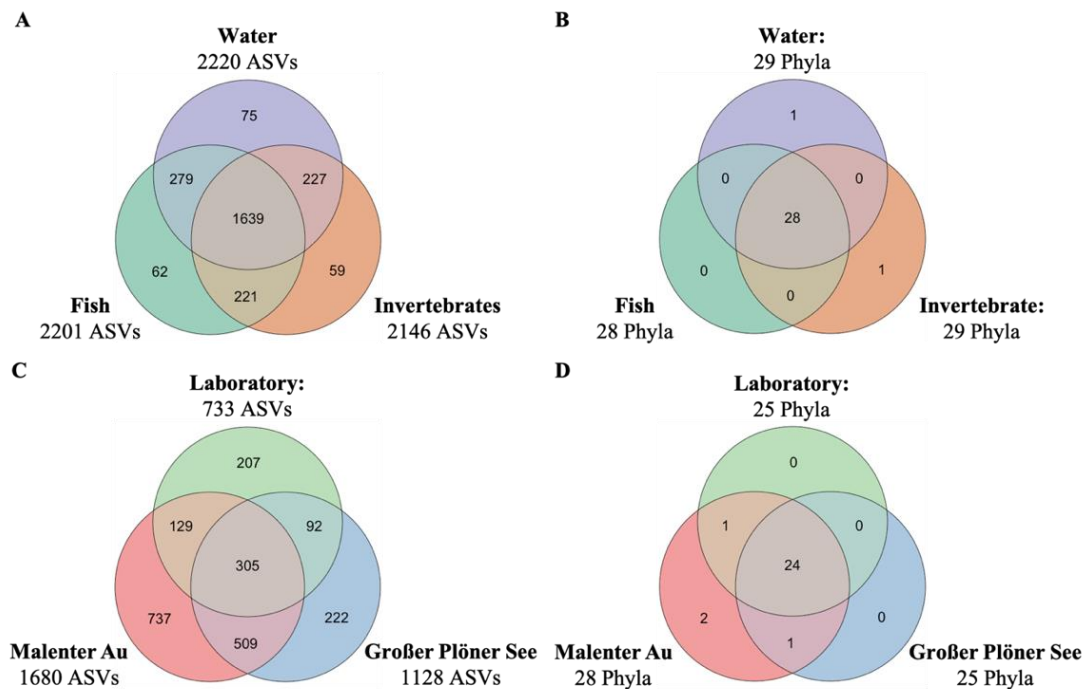
Figure 4.3: A) Shared and distinct ASVs across fish, water and invertebrate samples. B) Shared and distinct phyla across lake, fish, water and invertebrate samples. C) Fish shared and distinct ASVs across habitats of exposure, Malenter Au, Großer Plöner See and laboratory control. D) Fish shared and distinct phyla across habitats of exposure, Malenter Au, Großer Plöner See and laboratory control.

4.4.4 A comparison of the microbes in fish, invertebrates, and water

We identified an interaction between sample type and habitat of exposure across both experimental habitats and controlled laboratory settings (LMM, Shannon: $F_{4,432}$= 7.40, p < 0.001, PD: $F_{4,432}$= 5.32, p < 0.001, Gini-Simpson: $F_{4,432}$= 8.02, p < 0.001, Figure 4.4). Among the different habitats of exposure, we found microbial diversity was higher in invertebrate samples than fish samples from Großer Plöner See and Malenter Au (TukeyHSD, all diversity indexes: p < 0.01, Supplementary Table 4.7). Within the experimental habitats of exposure, microbial diversity did not differ between fish and

water samples (TukeyHSD, all diversity indexes: $p > 0.05$, Supplementary Table 4.7). In the control mesocosms, fish had higher microbial diversity than water samples (TukeyHSD, Shannon diversity and Gini-Simpson: $p < 0.01$, Supplementary Table 4.7). Among sample types, invertebrate microbial diversity did not differ between Großer Plöner See and Malenter Au, however, the invertebrates used to feed the laboratory fish harboured less diverse microbiomes than wild invertebrates from Großer Plöner See or Malenter Au (TukeyHSD, all diversity indexes: $p < 0.05$, Supplementary Table 4.7). Only Malenter Au and laboratory water samples differed in microbial diversity (TukeyHSD, Shannon diversity and Gini-Simpson: $p < 0.05$, Supplementary Table 4.7). Of greatest interest is that the microbial diversity of fish varied among all three habitats of exposure (TukeyHSD, all diversity indexes: $p < 0.05$, Supplementary Table 4.7).
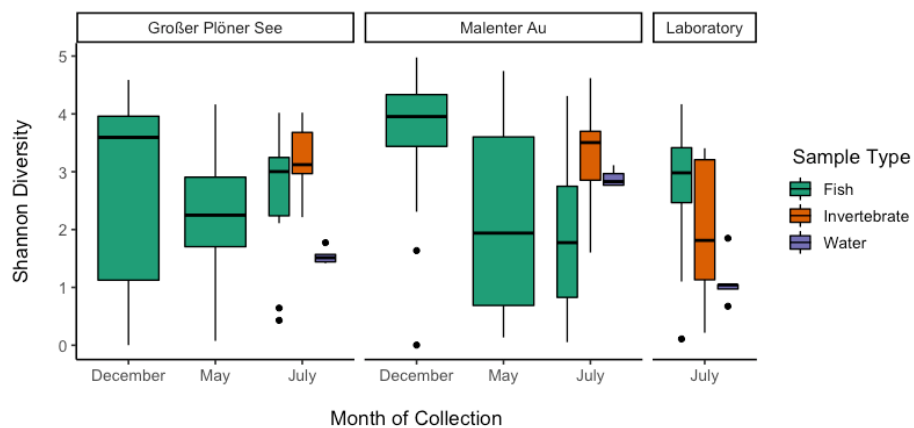
Figure 4.4: Boxplots show Shannon diversity significantly correlated with an interaction between sample type and habitat of exposure (LMM, Shannon: $F_{4,432}= 7.40$, $p < 0.001$). Shannon diversity is split by sampling months to reflect the random effect used in the statistical model.

Microbial community composition was best described by an interaction between sample type and habitat of exposure (PERMANOVA, Unweighted UniFrac: $F_{4,441} = 4.39$, $p < 0.001$, Weighted UniFrac: $F_{4,441}= 5.68$, $p < 0.001$, Bray-Curtis: $F_{4,441}= 4.80$, $p < 0.001$, Figure 4.5, Supplementary Table 4.8). Noteworthy, data dispersion may explain a part of the observed pattern (betadisper, all β-diversity metrics: $p < 0.001$, Supplementary Table

4.8). All pairwise interactions between sample types and locations were significant (Supplementary Table 4.9). Overall, it is obvious that each sample type is composed of its own microbiome, suggesting host-specific effects contribute to fish microbial community. We therefore focused on fish exposed to the different experimental habitats of exposure for the remainder of the analysis.
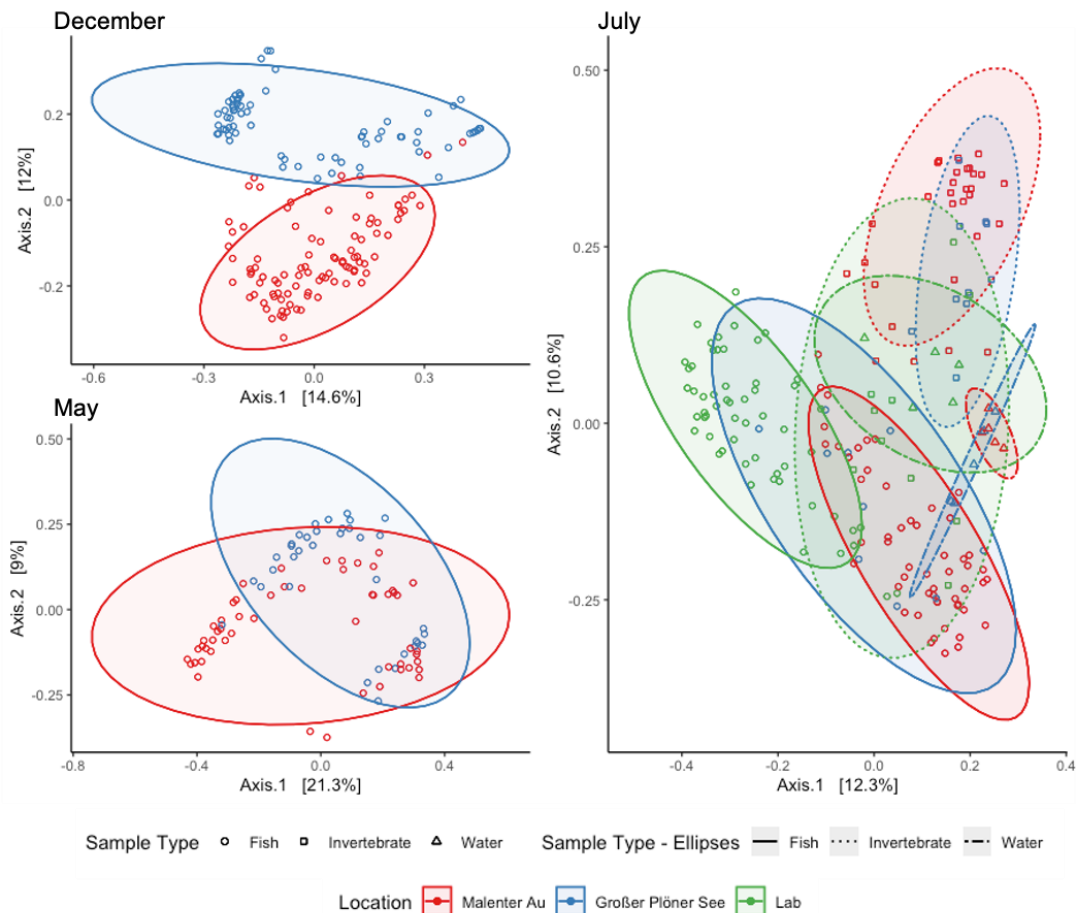
Figure 4.5: Principal Coordinates Analysis (PCoA) plot of the unweighted UniFrac distance microbial community of different habitats of exposure (colour) and sample types (shape) (PERMANOVA, $F_{1,230}$ = 5.03, p < 0.01). Plots are split by month of collection as sampling period was used as a block in the PERMANOVA.

4.4.5 Do fish in different habitats have different microbiomes?

To first describe the determinants of microbial diversity, we conducted a variance component analysis. We found that month of collection explained the largest component of the observed variation in microbial diversity (VCA, mean of all diversity index: 19.77

± 9.17% SD, Table 4.1). Mesocosm ID also explained a significant effect (VCA, 14.04 ± 6.76%, Table 4.1). Additionally, fish habitat of exposure, country of origin and family were small components of the observed microbial variation (Table 4.1).

Focusing on fish across both experimental habitats of exposure, we found microbial diversity correlated with an interaction between habitat of exposure and fish parasite load I_{pi}, with microbial diversity of Malenter Au fish being more negatively correlated with I_{pi} than Großer Plöner See fish (LMM, Shannon: $F_{1,184} = 8.82$, $p < 0.01$, PD: $F_{1,230} = 5.23$, $p < 0.05$, Gini-Simpson: $F_{1,209} = 5.07$, $p < 0.05$, Figure 4.6A). Furthermore, we found that fish origin was associated with microbial diversity, with German fish having a higher microbial diversity than Canadian fish for Shannon diversity index (LMM, $F_{1,32} = 4.49$, $p < 0.05$, Figure 4.6B). and Gini-Simpson index (LMM, $F_{1,32} = 4.80$ $p < 0.05$).

Table 4.1: Variance componence analysis explaining variability in fish microbial diversity due to month of collection, habitat of exposure, fish country of origin, ecotype, family group and mesocosm ID. Average of all indexes was calculated by calculating the mean of the combined variance components of the three indexes.

Factor	Average of all indexes		Shannon Diversity		Phylogenetic Diversity		Gini-Simpson Diversity	
	% Total	SD	% Total	SD	% Total	SD	% Total	SD
Total	-	-	100.00	1.56	100.00	5.47	100.00	0.33
Month of Collection	19.77	9.17	24.56	0.77	25.54	2.76	9.19	0.10
Habitat of Exposure	2.69	0.58	3.29	0.28	2.66	0.89	2.13	0.05
Country of Origin	5.95	2.23	6.67	0.40	3.44	1.01	7.72	0.09
Ecotype	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mesocosm ID	14.04	6.76	16.50	0.63	6.39	1.38	19.24	0.14
Family	0.27	0.23	0.29	0.08	0.04	0.11	0.49	0.02
Error	57.28	7.45	48.69	1.09	61.92	4.30	61.23	0.26
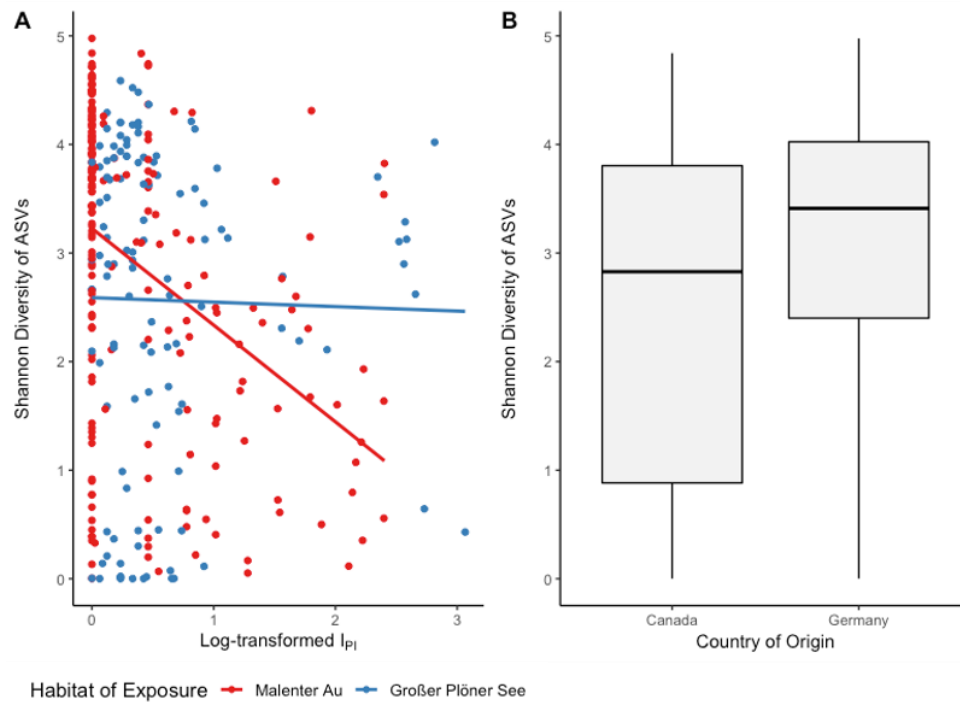
Figure 4.6: Shannon diversity was linked to A) an interaction between habitat of exposure and fish I_{pi} (LMM, Shannon: $F_{1,184} = 8.82$, $p < 0.01$), fish I_{pi} was log transformed for visualisation and B) fish country of origin (LMM, $F_{1,32} = 4.49$, $p < 0.05$).

Microbial community composition of fish was best described by an interaction between habitat of exposure and I_{pi} group (PERMANOVA, Unweighted UniFrac: $F_{2,319} = 1.85$, $p < 0.01$, Weighted UniFrac: $F_{2,319} = 2.18$, $p < 0.05$, Bray-Curtis: $F_{2,319} = 1.52$, $p < 0.05$, Supplementary Table 4.8). Specifically, within Malenter Au high I_{PI} infected fish showed a significantly different community composition to both low and medium I_{PI} infected fish (pairwise PERMANOVA, Unweighted UniFrac: $F_{1,101} = 7.53$, $p < 0.01$, Weighted UniFrac: $F_{1,101} = 6.09$, $p < 0.01$, Bray-Curtis: $F_{1,101} = 3.95$, $p < 0.01$, Supplementary Table 4.10). No significant pairwise interactions were observed within fish in Großer Plöner See (Supplementary Table 4.10). No ASVs were identified in significantly different abundance between habitats of exposure and I_{PI} pairs after FDR corrections. We found a second interaction between fish habitat of exposure and country of origin significantly correlated with microbial community (PERMANOVA, Unweighted UniFrac: $F_{1,319} = 4.04$, $p < 0.01$, Weighted UniFrac: $F_{1,319} = 4.91$, $p < 0.01$, Bray-Curtis: $F_{1,319} = 3.90$, $p < 0.01$, Figure 4.7, Supplementary Table 4.8). Within both

habitats of exposure, German and Canadian fish had significantly different microbial composition (pairwise PERMANOVA, all β-diversity metrics: $p < 0.01$, Supplementary Table 4.11). We also found that fish from the same country of origin had different microbial community composition across the two habitats of exposure (pairwise PERMANOVA, all β-diversity metrics: $p < 0.01$, Supplementary Table 4.11). These differences stemmed from 16 ASVs found in significantly different abundance between combined habitat of exposure and country groups (SIMPER, Supplementary Table 4.12). These ASVs belonged to the following phyla: Cyanobacteria, Firmicutes, Planctomycetes, Proteobacteria, Tenericutes, Verrucomicrobia. Here as well, data dispersion may have impact the observed results (betadisper, all β-diversity metrics: $p < 0.05$, Supplementary Table 4.8).
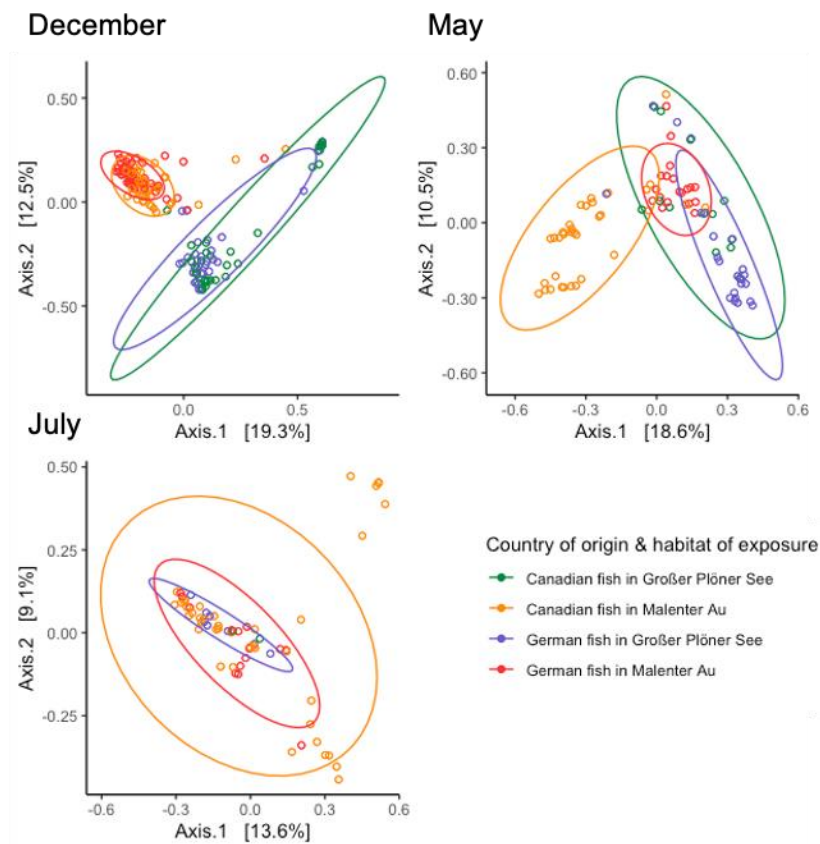
Figure 4.7: Principal Coordinates Analysis (PCoA) plot of Bray-Curtis distance of different country of origins and habitat of exposure (PERMANOVA, $F_{1,319} = 3.90$, $p < 0.01$). Plot are split by month of collection as it was used as a block in the PERMANOVA.

140

4.4.6 Microbial diversity and community in separated habitats of exposure

4.4.6.1 Fish exposed to Malenter Au river conditions.

Having determined that habitat of exposure correlated strongly with microbial diversity and composition, we investigated habitat-specific effects which may be otherwise hidden in a global analysis. Within the Malenter Au river, we found microbial diversity was associated with an interaction between fish country of origin and I_{PI}, with microbial diversity of German fish being more negatively correlated with I_{PI} than Canadian fish (LMM, Shannon: $F_{1,195}= 9.41$, p < 0.01 and Gini-Simpson: $F_{1,191}= 8.74$, p <0.01, Figure 4.8A). We found a second significant interaction between I_{PI} and sex of fish, where microbial diversity of females was more negatively correlated with I_{PI} than males (LMM, Shannon: $F_{1,192}= 6.32$, p < 0.05 and Gini-Simpson: $F_{1,193}= 13.81$, p < 0.001, Figure 4.8B). Interestingly, no specific ecotype effects were detected.
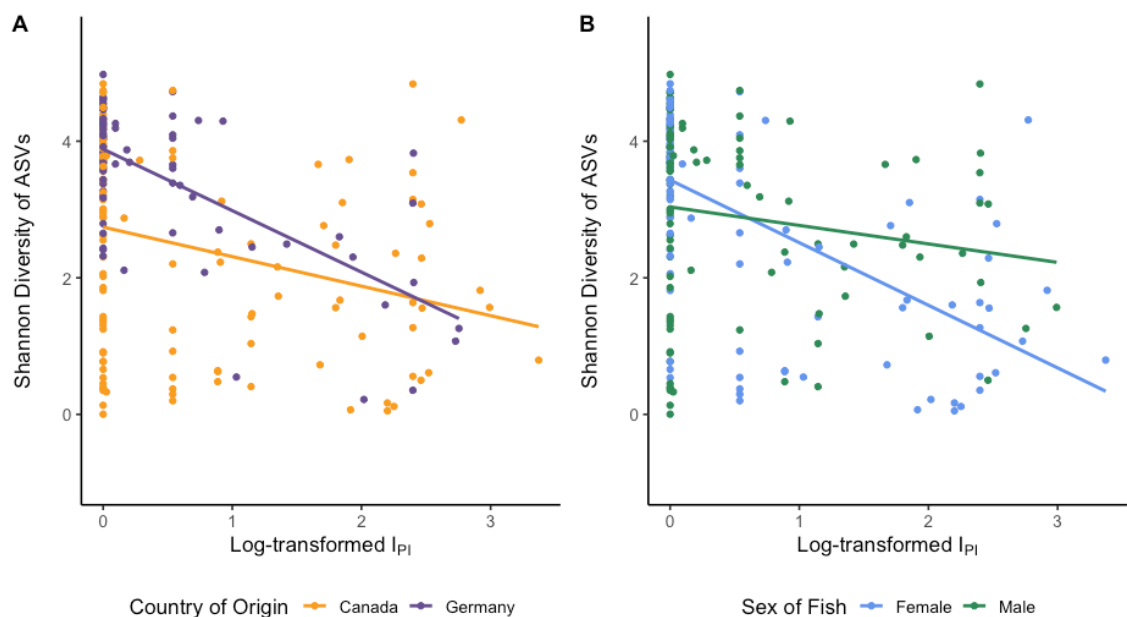
Figure 4.8: Within the Malenter Au habitat of exposure Shannon diversity is linked to A) an interaction between I_{PI} and country of origin (LMM, Shannon: $F_{1,195}= 9.41$, p < 0.01). B) an interaction between I_{PI} and sex of fish (LMM, Shannon: $F_{1,192}= 6.32$, p < 0.05). I_{PI} log-transformed for visual representation.

Within Malenter Au, microbial community showed significant variation with an interaction between fish country of origin and I_{PI} infection load, and we observed tighter

clusters in German fish, suggesting they have more similar microbiomes that Canadian individuals (PERMANOVA, Bray-Curtis: $F_{2,202} = 1.51$, $p < 0.05$, Figure 4.9, Supplementary Table 4.8). Specifically, microbial communities between German and Canadian fish varied when comparing all I_{PI} groups (pairwise PERMANOVA, $p < 0.05$, Table 4.2). Within each country of origin, fish with medium I_{PI} infection harboured significantly differently microbial communities to fish with a high I_{PI} infection (pairwise PERMANOVA, $p < 0.05$, Table 4.2). However, data dispersion may have impacted the observed results (betadisper, all β-diversity metrics: $p < 0.05$, Supplementary Table 4). We then performed a SIMPER analysis to identify the ASVs contributing to the observed difference. We detected 16 ASVs significantly contributing to the difference between country of origin and I_{PI} groups (Supplementary Table 4.13). Particularly *Carnobacterium viridans* was consistently found in higher abundance in Canadian fish than German fish, regardless of I_{PI} groups.
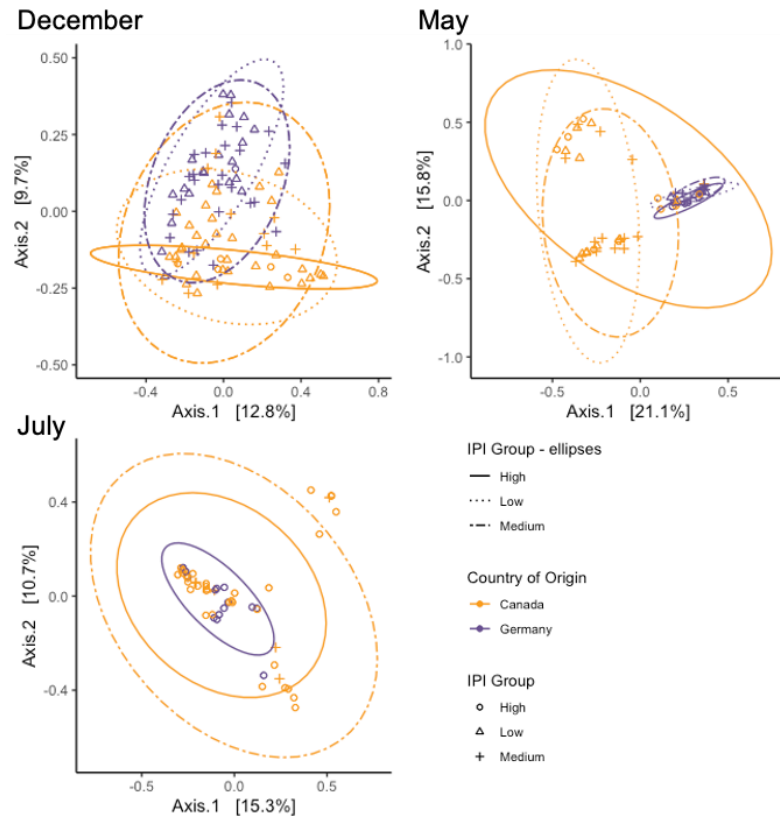
Figure 4.9: Within Malenter Au, Bray-Curtis diversity of fish showed significant variation with an interaction between country of origin and I_{PI} group (PERMANOVA, Bray-Curtis: $F_{2,202} = 1.51$, $p < 0.05$). Plot are split by month of collection as it was used as a block in the PERMANOVA.

Table 4.2: Selected Pairwise PERMANOVA results (pairwise.adonis2 function) from Malenter Au dataset for the interaction of country of origin and I_{PI} group using Bray-Curtis distance. Month of collection used as block. Significant results are highlighted in bold, Df denotes degrees of freedom.

	Df	Sums Of Sqs	Mean Sqs	F.Model	R2	Pr(>F)
German Fish						
Low I_{PI} - High I_{PI}	1,59	0.40	0.40	1.23	0.02	0.148
Medium I_{PI} - High I_{PI}	**1,54**	**1.62**	**1.62**	**4.55**	**0.08**	**0.023**
Low I_{PI} - Medium I_{PI}	1,44	1.51	1.51	4.09	0.09	0.094
Canadian Fish						
Low I_{PI} - High I_{PI}	1,89	1.41	1.41	3.43	0.04	0.098
Medium I_{PI} - High I_{PI}	**1,79**	**1.14**	**1.14**	**2.69**	**0.03**	**0.007**
Low I_{PI} - Medium I_{PI}	1,75	0.74	0.74	1.84	0.02	0.148
Low I_{PI} group						
German Fish - Canadian Fish	**1,67**	**1.18**	**1.18**	**3.16**	**0.05**	**0.001**
Medium I_{PI} group						
German Fish - Canadian Fish	**1,67**	**1.54**	**1.54**	**4.16**	**0.06**	**0.001**
High I_{PI} group						
German Fish - Canadian Fish	**1,66**	**0.60**	**0.60**	**1.43**	**0.02**	**0.045**

4.4.6.2 Fish exposed to Großer Plöner See lake conditions

None of the variables tested were significantly correlated with microbial diversity for fish exposed to the Großer Plöner See lake conditions. However, we found a number of effects for community composition. Once again, we found that country of origin was correlated with microbial community composition (PERMANOVA, Weighted UniFrac: $F_{1,116} = 3.74$, $p < 0.05$ and Bray-Curtis: $F_{1,116} = 2.61$, $p < 0.05$, Figure 4.10, Supplementary Table 4). We found no difference in dispersion of country groups within the Großer Plöner See exposure habitat (betadisper, $p > 0.05$, Supplementary Table 4). Interestingly, a single ASV, *Rickettsiella* species, was found in a significantly higher abundance in Canadian fish compared to German fish. Here as well, we found no effect of ecotype of origin.
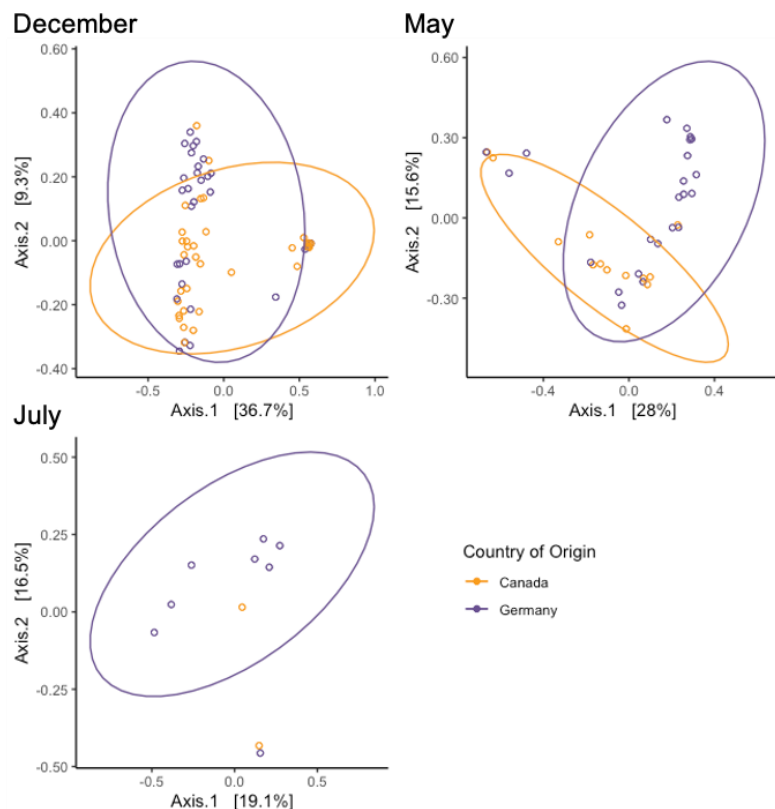
Figure 4.10: Country of origin was correlated with microbial community composition of fish within Großer Plöner See (PERMANOVA, Bray-Curtis: $F_{1,116} = 2.61$, $p < 0.05$). Plot are split by month of collection as it was used as a block in the PERMANOVA.

4.4.7 The influence of feeding ecology on the gut microbiome

For a subset of fish, we obtained stable isotope readings for carbon and nitrogen. Interestingly, Shannon diversity of river ecotype fish was negatively correlated with fish trophic position (LMM, Shannon: $F_{1,31}= 4.47$, $p < 0.05$, Figure 4.11A). Next, we found that an interaction between fish country of origin and proportional use of littoral carbon was associated with variation in phylogenetic diversity (LMM, PD: $F_{1,32}= 4.84$, $p < 0.05$, Figure 4.11B). Particularly, microbial phylogenetic diversity of Canadian fish decreased with proportion of littoral carbon, whilst German fish were positively correlated. We also found a second interaction between fish proportional use of littoral carbon and their habitat of exposure correlating with changes in phylogenetic diversity (LMM, PD: $F_{1,30}= 7.15$, $p < 0.05$, Figure 4.11C). Fish exposed to Malenter Au showed a weak positive correlation between microbial phylogenetic diversity and proportion of littoral carbon, whilst Großer Plöner See exhibited a negative correlation.
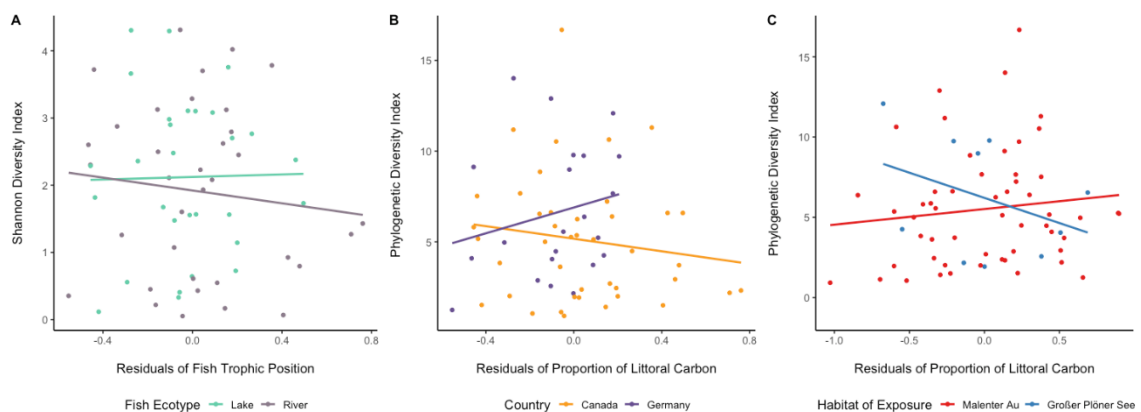
Figure 4.11: A) Shannon diversity correlated with fish ecotype and trophic position (LMM, Shannon: $F_{1,31}= 4.47$, $p < 0.05$). B) Phylogenetic diversity correlated with an interaction between country of origin and proportional use of littoral carbon (LMM, PD: $F_{1,32}= 4.84$, $p < 0.05$). C) Phylogenetic diversity correlated with an interaction between fish proportional use of littoral carbon and the habitat of exposure (LMM, PD: $F_{1,30}= 7.15$, $p < 0.05$).

A total of 388 ASVs had $> 0.01\%$ relative abundance and were tested for correlations with fish proportional use of littoral carbon and trophic position. After false discovery rate corrections, the relative abundance of 32 ASVs were associated with fish

trophic position (Figure 4.12). For the majority of ASVs, their relative abundance was positively correlated with fish trophic position. The relative abundance of 20 ASVs were associated with fish proportional use of littoral carbon, however after false discovery rate corrections, the number of significant ASVs was not significantly greater than the 5% expected due to false positives alone (p = 0.889).
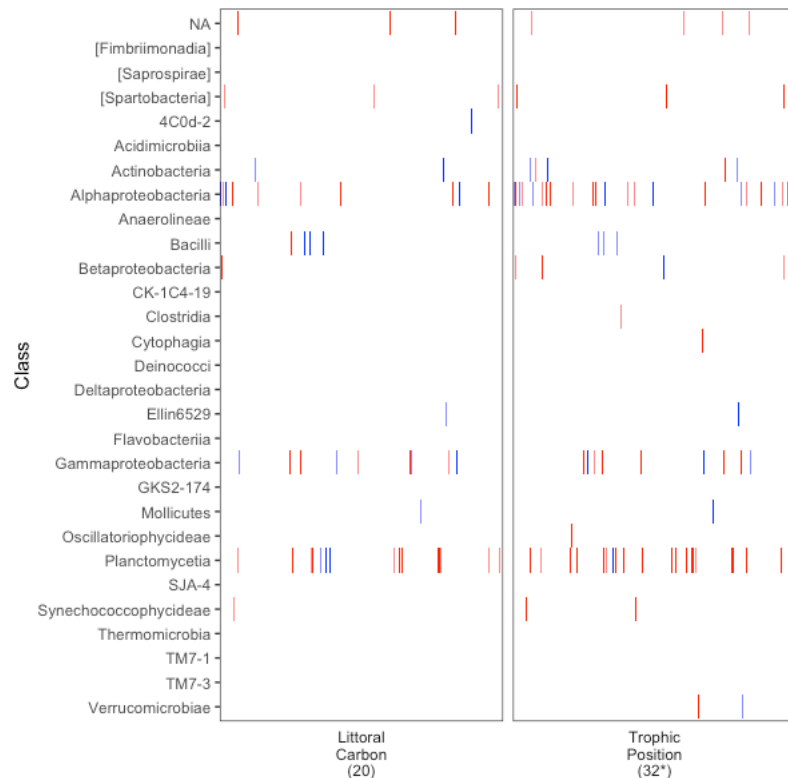
Figure 4.12: Heatmap showing the effects of diet on ASV relative abundance. Columns contain the models of diet effects, littoral carbon and trophic position for ASVs from a bacterial class in each row. Each row of the heatmap corresponds to a bacterial class. Vertical bars represent an ASV with a mean relative abundance >0.01%. Red bars represent ASVs whose relative abundance increases with the diet metric. The blue bars represent ASVs whose relative abundance decreased with the diet metric. Under the different diet metrics we indicate the number of ASVs for which an effect of the metric was observed. Asterisk indicates that the number of significant ASVs surpasses the expected 5% false positive rate.

4.4.8 Microbes and fish fitness

All models stated below show an interaction between ASVs of interest and habitat of exposure. See Table 4.3 for summary of models. Within Canadian fish, the presence of *Carnobacterium viridans*, which had a greater abundance in Malenter Au, correlated

with fish standard length (SL) (LM, SL: $F_{1,172}$= 4.23, p < 0.05, Figure 4.13A). *Serratia* species (*Serratia* species 1), which was found in higher abundance in Malenter Au, correlated with fish SL and splenosomatic index (SSI) (LM, SL: $F_{1,172}$= 10.16, p < 0.01, Figure 4.13B and SSI: $F_{1,172}$= 4.29, p < 0.05, Figure 4.13C). The remaining 5 ASVs identified for Canadian fish were not associated with standard length, body condition or splenosomatic index. In German fish, we found *Rickettsiella* species 1 in higher abundance in Großer Plöner See and was correlated with fish SL and body condition (CF) (LM, SL: $F_{1,140}$= 7.16, p < 0.01, Figure 4.13D and CF: $F_{1,140}$= 6.47, p < 0.05, Figure 4.13E). A second *Rickettsiella* species (*Rickettsiella* species 2), which had an increased abundance in Malenter Au, correlated with fish SL and CF (LM, SL: $F_{1,140}$= 15.50, p < 0.001, Figure 4.13F and CF: $F_{1,140}$= 9.08, p < 0.01, Figure 4.13G). The presence of *Luteolibacter* species 1, found in greater abundance in Malenter Au fish, correlated with fish SL and CF (LM, SL: $F_{1,140}$= 21.60, p < 0.001, Figure 4.13H and CF: $F_{1,140}$= 10.18, p < 0.01, Figure 4.13I). *Bacillus* species 1, higher in Großer Plöner See correlated with fish SL and CF (LM, SL: $F_{1,140}$= 6.56, p < 0.05, Figure 4.13J and CF: $F_{1,140}$= 7.97, p < 0.01, Figure 4.13K). Next, we found an ASV belonging to the class CK-1C4-19, which was found in higher abundance in Malenter Au fish, correlated with fish SL (LM, SL: $F_{1,140}$= 2.41, p < 0.05, Figure 4.13L). Gemmataceae family 1, again found in higher abundance in Malenter Au fish, correlated with all three fitness traits (LM, SL: $F_{1,140}$= 7.93, p < 0.01, Figure 4.13M, CF: $F_{1,140}$= 5.75, p < 0.05, Figure 4.13N, and SSI: $F_{1,140}$= 4.16, p < 0.05, Figure 4.13O). Finally, the presence of a *Rhodobacter* species (*Rhodobacter* species 1), which was in higher abundance in Großer Plöner See fish, correlated with CF and SSI (LM, CF: $F_{1,140}$= 4.85, p < 0.05, Figure 4.13P and SSI: $F_{1,140}$= 8.18, p < 0.01, Figure 4.13Q).

Table 4.3: Summary table of linear models using habitat of exposure and individual ASVs to explain fish fitness traits, standard length (SL), body condition (CF) and splenosomatic index (SSI), split into Canadian fish and German fish. Found in greater abundance refers to the habitat in which the ASV is found in higher abundance (MAU – Malenter Au, GPS – Großer Plöner See).

			Fish Fitness Trait					
ASV	*Found in greater abundance in:*	*Df*	*Standard Length*		*Body Condition*		*Splenosomatic index*	
			F value	*P value*	*F value*	*P value*	*F value*	*P value*
Canadian Fish:								
Carnobacterium viridans	MAU	1,172	**15.50**	**< 0.001**	3.68	0.057	0.21	0.644
Serratia species 1	MAU	1,172	**4.29**	**0.040**	1.80	0.182	**10.16**	**0.002**
German Fish:								
Rickettsiella species 1	GPS	1,140	**7.16**	**0.008**	**6.47**	**0.012**	2.43	0.121
Rickettsiella species 2	MAU	1,140	**15.50**	**< 0.001**	**9.08**	**0.003**	1.17	0.281
Luteolibacter species 1	MAU	1,140	**21.60**	**< 0.001**	**10.18**	**0.002**	1.91	0.170
Bacillus species 1	GPS	1,140	**6.56**	**0.011**	**7.98**	**0.005**	0.09	0.759
CK-1C4-19 class	MAU	1,140	**2.41**	**0.123**	3.60	0.060	0.17	0.679
Gemmataceae family 1	MAU	1,140	**7.93**	**0.006**	**5.75**	**0.018**	**4.16**	**0.043**
Rhodobacter species 1	GPS	1,140	0.17	0.678	**4.85**	**0.029**	**8.18**	**0.005**
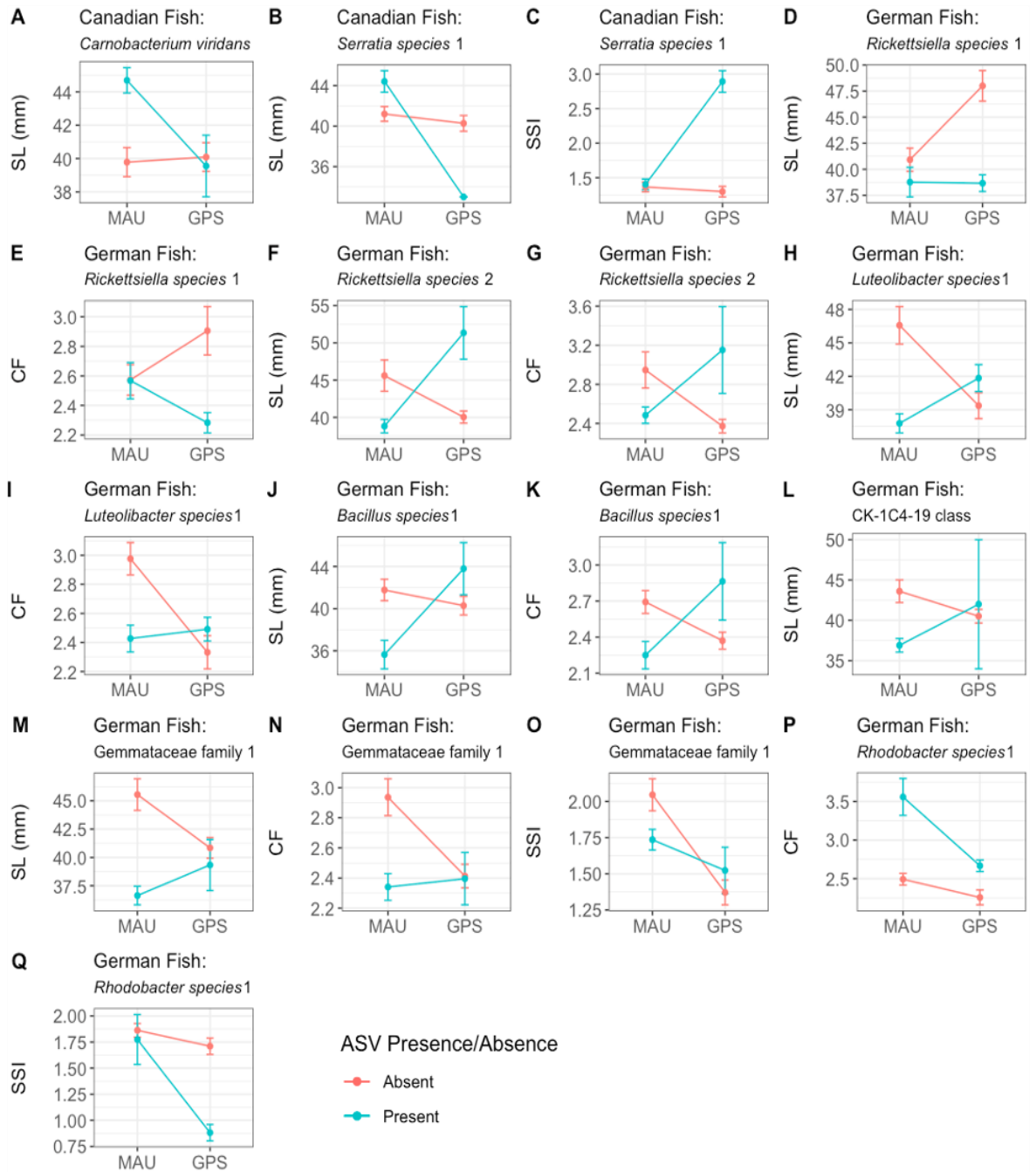
Figure 4.13: Norm of reaction showing influence of presence/absence of ASVs on different fish fitness traits: standard length (SL), body condition (CF) and splenosomatic index (SSI) within experimental habitats of exposure, Malenter Au (MAU) and Großer Plöner See (GPS) separated by country of origin, Canadian and German fish.

4.5 Discussion

The influence host (genetic) and environmental factors have on the host microbiome are difficult to disentangle in natural populations (Rennison *et al.*, 2019a). In this study, we attempted to address this knowledge gap, using Canadian and German parapatric lake-river three-spined stickleback pairs, translocated into new lake and river habitats in Germany. Firstly, we found evidence of host contribution, as i) the microbiome of the three-spined stickleback was distinct from that of their food source (invertebrates) and the surrounding water; and ii) fish country of origin was repeatedly associated with variation gut microbiome even in shared habitats of exposure. We showed, on the other hand, host environment correlated with changes in the gut microbiome, through temporal variation (highlighted in the VCA results) and habitat-specific microbiomes. We revealed genotype-by-environment interactions through correlations between the gut microbiome and parasite infection, as well as diet. Finally, we identified signals of host local adaptation (at the country level) through fish survival and exploring the influence of individual ASVs on fish fitness traits. Altogether, our results confirm that host's genetics and their environment correlate strongly with changes in the gut microbiome in the wild, with complex host-parasite-microbe and host-microbe-diet interactions occurring.

The influence of stickleback genetics on microbial diversity and composition was identified repeatedly throughout our study. Firstly, we can confirm fish gut microbiota are not simply a subset of the microbes of their prey and surrounding water as we found 62 fish-specific ASVs (belonging to a mix of phyla but high numbers of: Planctomycetes Proteobacteria and Firmicutes), as well as significant variation in the microbial diversity and composition between the invertebrates and water samples across exposure habitats. This could be driven by host environmental filtering, which has been shown to regulate the microbial community in freshwater fish (Bolnick *et al.*, 2014b, Yan *et al.*, 2016).

150

Additionally, some organisms can actively filter environmental microbes, allowing certain species to colonise whilst avoiding others, such as in the case of the Hawaiian bobtail squid, *Euprymna scolopes,* and *Vibrio fischeri* (Nyholm and McFall-Ngai, 2004, Ohbayashi *et al.*, 2015). Secondly, we found fish country of origin repeatedly correlated with variation in the gut microbiome, both as main effects and as interactions with habitat of exposure and fish parasite load, I_{PI}. One potential explanation relates to a combination of drift, demographic effect and local selection pressures in the ancestral environment which have led to changes in fish genetic architecture (Chain *et al.*, 2014, Feulner *et al.*, 2015), resulting in country-specific variation in the fish microbiome, even when placed in the same exposure habitat. Interestingly, we did not find strong evidence of microbial variation between lake and river ecotype pairs. This could suggest that the genetic difference among pairs is still not strong enough to overcome the influence of the local environment on the host microbiome, potentially because lake-river ecotypes differentiated much more recently than Canadian and German lineages (Chain *et al.*, 2014, Feulner *et al.*, 2015). Genetic variation in sticklebacks is observed globally (Feulner *et al.*, 2015, Marques *et al.*, 2016, Ravinet *et al.*, 2013). However fish from the northern Pacific show increased standing genetic variation compared to fish from the Atlantic basin, with a large number of alleles being lost as sticklebacks expanded out of the Pacific (Fang *et al.*, 2020). We also found fish sex, which is genetically determined, interacted with I_{PI} in the Malenter Au habitat, with females' microbial diversity being more negatively correlated with parasite infection than males. Stickleback sex has previously been linked to changes in the gut microbiome, revealing a genetic component mediating microbial diversity (Bolnick *et al.*, 2014c). Sex-specific microbial variation may be driven by sex-specific immune responses and hormones – a common trait of fish metabolism (Koren *et al.*, 2012, Markle *et al.*, 2013).

This study also highlighted environmental influences on the stickleback gut microbiome. We found 20% of the total variance in microbial diversity was explained by the month the fish were sampled. This temporal effect is likely driven by seasonal changes in abiotic variables and ecological communities (Friberg *et al.*, 2019, Uren Webster *et al.*, 2020). Previous studies on other wild organisms, such as mice and frogs, have highlighted seasonal changed are a strong driving factor behind microbial variation (Maurice *et al.*, 2015, Xu *et al.*, 2020). An alternative explanation may be that we are seeing a successional change in the microbiome with fish age, as diversity and composition vary greatly throughout fish development (Stephens *et al.*, 2016, Yan *et al.*, 2016). This could be explored through controlled experiments in the laboratory by sampling fish of the same genotype and age housed in different environmental conditions.

In our study, habitat of exposure repeatedly correlated with variation in the gut microbiome. Firstly, we found a large number of ASVs that were habitat of exposure specific, with Malenter Au harbouring two unique core ASVs, *Propionibacterium acnes* and a *Rickettsiella* species. *Propionibacterium acnes* is a common bacteria found in the intestine of freshwater and marine fish (Austin, 2006, Green *et al.*, 2013). Interestingly, members of the genus *Rickettsiella* are known intracellular bacterial pathogens of arthropods which can also be pathogenic if transferred to vertebrates, with some species causing typhus and Rocky Mountain spotted fever (Cordaux *et al.*, 2007, Perlman *et al.*, 2006). We also identify two phyla that were unique to Malenter Au fish, Synergistetes, which is often found in the microbiota of animals (Godon *et al.*, 2005) and SR1 which has previously been found in high-temperature marine environments and fresh-water lakes (Davis *et al.*, 2009). Fish held in Großer Plöner See had three core ASVs, a *Synechococcus* species, a *Rickettsiella* species and an ASV belonging to the Isosphaeraceae family. The genus *Synechococcus* belongs to the phylum cyanobacteria

and are classed as key components of freshwater picophytoplankton (Callieri, 2008). Secondly, we showed that fish from the same country harboured different microbiomes across the two exposure habitats. A similar study on translocated salmon showed a near-complete turn over in gut microbiome composition pre and post-translocation, which was correlated with environmental factors, whilst host genetics had little impact (Uren Webster *et al.*, 2020). This suggests the relative influence these factors have on the microbiome may vary across species as stickleback did not show a complete turnover of their microbiome community. Altogether, these results show clear environmental influence on the gut microbiome. While we did not measure temperature or pH for instance, it is likely that lake and stream abiotic conditions vary. We measured however the biotic components of the environment.

Thanks to our field experimental exposure, we have been able to evaluate the link between biotic environmental factors, in particular parasite infection and diet, and fish microbiome. Whether infection or predation, both showed genotype-by-environment interactions correlated with microbial diversity and composition. For instance, fish parasite load (I_{PI}) was associated with changes in the gut microbiome through interactions with habitat of exposure, country of origin and sex. A large body of research has focused on stickleback host-parasite interactions, showing sex and ecotype-specific infections, with genetic variation in immune genes driving variation in parasite resistance (Brunner *et al.*, 2017, Eizaguirre *et al.*, 2012a, Eizaguirre *et al.*, 2012b, Eizaguirre *et al.*, 2009b, Feulner *et al.*, 2015, Lenz *et al.*, 2013). Specifically, in the German system, fish in Großer Plöner See are exposed to a greater load and diversity of parasites than Malenter Au fish (Eizaguirre *et al.*, 2011, Kalbe *et al.*, 2002). Parasite communities play a role in driving the local adaptation of a host (Eizaguirre *et al.*, 2012b, Feulner *et al.*, 2015), but their relationship with a host's microbiome is often overlooked (Leung *et al.*, 2018). Host-

parasite-microbe interactions in stickleback remain relatively unexplored, with studies generally focusing on single infections within controlled conditions (Ling *et al.*, 2020). Parasite infection can change the physiology of a host, for example, helminth infections can alter the mucus production and epithelial cell turnover in the gastrointestinal tract of a host (Hasnain *et al.*, 2013, Kim and Khan, 2013, Tsubokawa *et al.*, 2015). These parasite-mediated changes impact the intestinal microbiome as the mucus layer is used by bacteria both as a food source and somewhere to live (Leung *et al.*, 2018). Of interest, two *Rickettsiella* species, that are potential intracellular bacterial pathogens, were consistently found in higher abundance in fish with low parasite load in comparison to highly infected groups (Cordaux *et al.*, 2007, Perlman *et al.*, 2006). We also identified a *Luteolibacter* species which has previously been found in the intestine of killifish (*Phalloceros caudimaculatus*) in highly polluted streams (Nolorbe-Payahua *et al.*, 2020), which is repeatedly found in higher abundance in low and medium infection groups than high infection groups, from both countries of origin. This suggests a negative interaction between parasite load and microbe infection, possibly as a result of immune responses preventing microbes from colonizing or residing within the host. Further laboratory experiments could test this direct link, further highlighting how field and laboratory studies can complement each other to identify ecologically relevant species interactions.

Additionally, we observed diet-microbe interactions where microbial diversity was positively correlated with the proportion of littoral carbon in a fish diet exposed to Malenter Au (river habitat). This correlation was however reversed in fish exposed to Großer Plöner See conditions. This indicated that lower microbial diversity was associated with fish consuming more benthic food sources, such as invertebrates in Großer Plöner See, but with more pelagic food sources such as zooplankton in Malenter Au. A similar pattern is seen between German and Canadian fish, where German fish

showed increased microbial diversity with more benthic food sources, whilst Canadian fish had decreased microbial diversity. Microbial diversity of lake ecotypes increased as they feed at higher trophic levels, whilst river ecotypes that feed at higher trophic levels had a reduced microbial diversity. These results show genotype-by-environment interactions and should be explored further in controlled laboratory settings to determine the causality, it is likely linked however with the lack of favourite prey items accessible in the mesocosms. Furthermore, the relative abundance of several individual ASVs were significantly associated with fish trophic position. These ASVs predominantly belonged to two bacterial classes. Firstly, Planctomycetia are common decomposers found in aquatic systems that feed on algae, and are known to assist isopod in digestion (Aires *et al.*, 2018). Secondly, Alphaproteobacteria are often found in wastewater (Kragelund *et al.*, 2006). As Planctomycetia is associated with invertebrates it is logical that we find it linked to trophic position of sticklebacks as they consume a wide range of invertebrates, often showing interindividual variation leading to varied stable isotope signals (Bolnick *et al.*, 2014b). Diet has been repeatedly linked to changes within a host microbiome across a range of wild organisms (Baldo *et al*., 2015, Brice *et al*., 2019, Friberg *et al*., 2019, Youngblut *et al*., 2019), these previous findings combined with the results demonstrated here show that the impact of diet is relevant across species and should be taken into consideration whenever studying host microbiomes.

When exposed to experimental habitats, Canadian fish in Großer Plöner See had a reduced likelihood of survival compared to Canadian fish in Malenter Au and German fish in either habitat. A reduction in the survival of a foreign genotype, Canadian fish, in comparison to a local genotype, German fish, in the same habitat is a signal of local adaptation (Kawecki and Ebert, 2004). However, it is important to highlight that this led

to low sample sizes from the Großer Plöner See mesocosms in the July time point which may impact some of our findings.

Whilst bacteria can be beneficial to a host (Alberdi *et al.*, 2016, Koskella *et al.*, 2017, McFall-Ngai *et al.*, 2013), they are also known to cross the parasite-mutualist continuum, becoming pathogenic in the absence of more virulent pathogens or in stressful environments (Chamberlain *et al.*, 2014, King *et al.*, 2016). It is possible that the presence or absence of certain bacteria could influence host fitness (Gould *et al.*, 2018, Shu *et al.*, 2018). Among the 16 ASVs that varied in abundance between the exposure habitats, nine were correlated with fish fitness. The presence of five ASVs had positive effects on fish fitness within the habitat where the ASV was found in higher abundance. The remaining four ASVs had negative or no effect on fish fitness. Positive effects of ASV presence included increased fish size, body condition and reduced splenosomatic index. Increased spleen size has been linked with costly immunological activation so a lower splenosomatic index and a higher body condition suggests fish are fitter (Kalbe *et al.*, 2009). These positive correlations suggest increased fitness of fish in the presence of individual ASVs and are to be expected if the evolution of local adaptation in the host also correlates with the evolution of host-specific microbiomes.

Our results are not limited simply to the three-spined stickleback or fish but can be used to infer the impact of host genetics and the environment on microbiomes across species, as the patterns observed in this study have also been shown in other species such as mice and amphibians (Knowles et al., 2019, Leung et al., 2018, Ling et al., 2020, Maurice et al., 2015). We show wild host-microbe interactions are influenced by a combination of host genetics and environmental pressures such as parasite infection and diet. The complex interactions between host-parasite-prey and microbes showed that exploring wild host microbiome in relation to parasites and diet is crucial to improve our

understanding of species' evolution. Field experiments also serve to validate laboratory-based findings as such complexity cannot be easily simulated in controlled conditions.

Ethics statement

All animal experiments described were approved by the Ministry of Nature, Environment and Country Development, Schleswig Holstein, Germany. Permits were granted by Canadian and German governmental institutions for all steps from catching fish in Canada and bringing them to Germany.

Chapter 5: Philopatric units and parasite infection determine microbial diversity and composition of loggerhead sea turtles.

Adrienne E Kerley[1], Leila Fouda[1], Doko-Miles J. Thorburn[1], Thomas Reischig[2], Albert Taxonera[3], Silvana Monteiro Roque [4], Leno Dos Passos[5], Lino Duarte[5], Lee Henry[1], China Hanson[6], Christophe Eizaguirre[1]

[1] Queen Mary University of London, School of Biological and Chemical Sciences, Mile end road, E1 4NS, London, UK

[2] Turtle foundation, An der Eiche 7a, 50678 Cologne, Germany.

[3] Associação Projeto Biodiversidade, Mercado Municipal 22, Santa Maria 4111, Ilha do Sal, Cabo Verde

[4] Projeto Vitó Porto Novo, Porto Novo, Ilha do Santo Antão, Cabo Verde

[5] Foundation Maio Biodiversity, Cidade de Porto Inglês, Ilha do Maio, Cabo Verde

[6] Microbiology@UCL, University College London, London WC1E 6BT, UK

5.1 Abstract

Host microbiomes are determined by a continuum of variables, ranging from the sole impact of the environment to the dual effect of environmental and host genetics. Under what conditions one is more important than the other is still highly debated, particularly in wild populations. Here, we focused on the loggerhead sea turtles, *Caretta*, that nest on different islands of the Cabo Verde archipelago, but which are known to spend the vast majority of their lifetime in shared feeding grounds. We investigated the effects of philopatric nesting groups (genetic effect) on different islands, parasite infection (environment) and turtle size (combined environment x genetics) on host internal cloacal microbiome. Proteobacteria, Bacteroidetes, Lentisphaerae, Fusobacteria and Firmicutes were the most common phyla detected. We found that the nesting island of a turtle, its infection status with the leech *Ozobranchus margoi*, as well as their interaction, were the best descriptors of microbial diversity and community composition, suggesting both environmental and genetic factors influence the cloacal microbiome. Specifically, turtles nesting on Santo Antão, the most genetically and geographically distinct nesting group sampled, carried distinct microbial communities compared to turtles nesting on all other islands. This pattern was driven by a greater abundance of specific amplicon sequence variants (ASVs) belonging to Lentisphaerae, Bacteroidetes, and Firmicutes. Our study shows that cloacal microbiome diversity and composition are influenced by genotype-by-environment interactions, which reflects local adaptation and indicates that despite living in a shared environment, the locally adapted turtle nesting groups harboured nesting group-specific microbiomes.

5.2 Introduction

Over the last decade, research has increased our understanding of the determinants of microbiome diversity and structure, as well as the role they play for a host's metabolism (Lee and Mazmanian, 2010, McFall-Ngai *et al.*, 2013). Particularly, microbiomes influence immune function and nutrient uptake, both contributing to the host's body condition and ultimately reproductive fitness (Alberdi *et al.*, 2016, Bouskra *et al.*, 2008, Lawley and Walker, 2013, Macpherson and Harris, 2004, McFall-Ngai *et al.*, 2013, Suchodolski, 2011).

Determining what drives variation in microbiome diversity and structure is complex, particularly in wild populations. To overcome challenges inherent to field studies, microbial research has often focused on humans, mice, and rats as they offer molecular resources which are still unmatched in other systems. However, studies have mostly been held under controlled conditions to account for unwanted environmental variation (Kostic *et al.*, 2013, Ley *et al.*, 2008a, Ley *et al.*, 2008b). Yet, individual hosts do not exist in isolation, and are part of an ecosystem that varies in ecological conditions. They also interact with conspecifics, parasites and prey, all of which are known to alter internal microbiomes (McFall-Ngai *et al.*, 2013). Given what we have learned from those controlled lab-studies, it is now time to expand to a more diverse range of wild species, and verify the causes and consequences of microbiome diversity and structure in natural settings (Hird, 2017). Studies focusing on wild organisms have increased in recent years exploring the wild microbiome of mice (Davidson et al., 2020, Raulo et al., 2021, Schmidt et al., 2019), amphibians (Jervis *et al*., 2021), primates (Lee et al., 2021, Tung et al., 2015) and fish (Dulski *et al*., 2020, Minich *et al*., 2020) but more studies with a focus on cryptic and endangered species and the potential insights into their health that can come from understanding their microbiome are needed.

A host's microbial composition is influenced by external environmental factors, as well as their genetic make-up shaping their internal physiology and metabolism (Spor *et al.*, 2011). Temperature, pH, pollutant exposure, and salinity are examples of external variables that alter a host's microbiome (Chiu *et al.*, 2020, Krause *et al.*, 2012, Lozupone and Knight, 2007, Meng *et al.*, 2018, Sullam *et al.*, 2012). Host diet also strongly correlates with microbial diversity and composition (Bloodgood *et al.*, 2020, Bolnick *et al.*, 2014b, David *et al.*, 2014). Herbivorous individuals, for example fish, harbour more cellulose-degrading bacteria such as *Clostridium, Citrobacter* and *Leptotrichia* compared to their omnivorous and carnivorous counterparts (Liu *et al.*, 2016). Interestingly, hosts' diets are not independent of other biotic interactions such as parasite infection since during foraging and feeding, hosts are exposed to diverse parasites (Schmid-Hempel, 2011). Once infected, genetic factors associated with the response to parasites even lead to different microbiomes compared to non-infected individuals (Ling *et al.*, 2020). This effect is well shown in the model fish system *Gasterosteus aculeatus*, that, once infected with the trophically-transmitted parasite *Schistocephalus solidus,* showed different fish families carried different microbial communities (Ling *et al.*, 2020). Such differences among families were, however, not detectable before the infection (Ling *et al.*, 2020). Other host genetic effects on host-parasite-microbiome interactions have been detected in the relative abundance of several microbial orders, such as Planctomycetales and Campylobacterales between male and female three-spined stickleback (Ling *et al.*, 2020). Indeed, prior to infection, males and females harboured distinct microbiomes, but this difference disappeared upon infection (Ling *et al.*, 2020). This is because the activation of the immune system of the host upon infection will change the host-microbiome interactions (Leung *et al.*, 2018), and also because parasites can also directly interact with the host microbiome, competing, for instance, for space (Dheilly *et al.*, 2015). Overall, it

is evident that host-parasite-microbiome interactions are becoming a prime focus in microbiology research, particularly so because certain elements of the microbiota can also protect the host against infection (Ford and King, 2016, Holm *et al.*, 2015, Jaenike *et al.*, 2010, White *et al.*, 2018).

Because host-parasite and host-microbiome interactions can be context-dependent, it is not surprising that local adaptation has been proposed as an evolutionary outcome of host-parasite-microbiome interactions (Brucker and Bordenstein, 2012, McFall-Ngai *et al.*, 2013, Rudman *et al.*, 2019). Local adaptation is detected when a resident population has increased fitness in their local environment compared to foreign migrants (Kawecki and Ebert, 2004, Savolainen *et al.*, 2013). Focusing on microbiomes, local adaptation can be observed when a locally adapted host shows population-specific microbiomes even when the environmental conditions are similar. Because host microbiomes may facilitate the evolution of host local adaptation across geographical scales, studying one species in a single location may only result in a partial understanding of a host microbial diversity and community (Rennison *et al.*, 2019a).

Whilst an impossible laboratory model organism, sea turtles make a good study system to test the influence of genetic and environmental factors in a natural system. Sea turtles are philopatric, returning to their place of birth to breed and deposit their eggs, and also have a wide distribution (Bowen *et al.*, 2004). Such natal behaviour reduces gene flow among nesting aggregations (Baltazar-Soares *et al.*, 2020, Shamblin *et al.*, 2014), and even among nesting groups within nesting aggregations (Baltazar-Soares *et al.*, 2020, Stiebens *et al.*, 2013b). In certain parts of the world, turtles show nesting group-specific feeding ecology (Cameron *et al.*, 2019), parasite infection (Lockley *et al.*, 2020) and immune gene diversity (Stiebens *et al.*, 2013a). All these findings suggest nesting groups are locally adapted to their nesting sites, even though turtles spend the majority of their

lifetime in a shared oceanic feeding environment, away from their nesting sites. Noteworthy, when it comes to microbial research, studies are biased towards mammals, fish and other model organisms, leaving important taxa unexplored, making sea turtles a very interesting yet understudied organism (Sullam *et al.*, 2012). Additionally, six out the seven sea turtle species are classified as "vulnerable" or higher on the IUCN red list (Casale and Tucker, 2017), therefore it is important to have a deep understanding of their physiology and biology in order to better protect them. Previous studies have highlighted the importance the host microbiome plays in the health, physiology, and behaviour of animals (Alberdi *et al.*, 2016, Koskella *et al.*, 2017, McFall-Ngai *et al.*, 2013) so it is crucial to improve our understanding of these cryptic, vulnerable organisms in order to aid in their conservation.

To date, most studies on the sea turtle microbiome have focused on individuals held in rehabilitation centres, which prevents testing for how turtles' local adaptation correlates with microbial patterns of diversity and community composition (Abdelrhman *et al.*, 2016, Ahasan *et al.*, 2018, Biagi *et al.*, 2019, Bloodgood *et al.*, 2020). These captive individuals are sick or in recovery states, so their microbiomes may differ from healthy, wild, individuals. Of the limited studies on the wild microbiome of sea turtles, Scheelings *et al.* (2020) found higher microbial diversity in the gut microbiome of loggerhead sea turtles from Australia compared to those nesting in the USA. They suggested variation between their microbial communities showed that environmental and genetic factors can influence turtles' microbiomes. Not only was the sample size of this study small (N_{USA}= 6, N_{Aus}= 18), but the scale of genetic differentiation was too large to test for local adaptation given the impossible gene flow between those rookeries and the lack of an overlapping feeding ground. As such, the patterns of microbial diversity and structure remain unclear among turtles studied at an ecologically-relevant scale. Noteworthy, the

size of loggerhead turtles correlated with changes in their microbial community structure with larger, older individuals showing less variation in community structure and small young individuals showing a wider spread of microbial communities (Biagi *et al.*, 2019).

Here, we focused on loggerhead sea turtles (*Caretta caretta*) nesting on islands in the Cabo Verde archipelago. Each island supports a unique nesting group, locally adapted in terms of feeding (Cameron *et al.*, 2019), parasite infection (Lockley *et al.*, 2020) and immune gene diversity (Stiebens *et al.*, 2013a), all of which have the potential to interact with the diversity and community structure of the microbiome (Dheilly, 2014, Ford and King, 2016, Spor *et al.*, 2011). Whilst loggerhead sea turtles can be infected with several different parasites, in this rookery, the leech *Ozobranchus margoi* correlates with feeding ecology, whereby oceanic turtles showed increased infection prevalence than their neritic counterparts (Lockley *et al.*, 2020). Furthermore, the infection by this ectoparasite has been linked with the evolution of size-dependent changes in reproductive strategy (Lockley *et al.*, 2020). There is limited knowledge about the life cycle and biology of these leech parasites, such as whether they can survive separate from the turtle or how they are transmitted between individuals, however it is known that they are capable of completing their entire life cycle on the turtle (McGowin *et al.*, 2011). In association with island-specific environmental conditions, parasite-mediated selection for local adaptation acting upon nesting groups could further result in nesting-group specific microbiomes. Focusing on turtles nesting on four islands, we tested whether microbial diversity and structure (1) vary among nesting groups to investigate microbial composition within a cryptic and vulnerable species potentially identifying local adaptation to nesting islands, (2) differ with infection by the leech *Ozobranchus margoi*, and (3) correlate with turtle size. Because certain storage methods may affect the microbe obtained from a sample (Choo *et al.*, 2015, Kim *et al.*, 2017a, Song *et al.*, 2016), we also tested (4) whether cloacal

microbial diversity and composition varied with sample storage methods. We used cloacal swab samples and next-generation sequencing to identify microbial diversity and composition.

5.3 Methods

5.3.1 Sample collection

Female turtles were sampled during the nesting season in 2019 on four islands of the Cabo Verde archipelago: Boa Vista, Maio, Sal, and Santo Antão ($T_{urtles} = 135$, Figure 5.1). Cloacal swab samples (Isohelix, SK-3S) were collected from female loggerhead sea turtles during oviposition. Swabs were inserted into the dilated cloaca and rolled gently on the cloacal wall without contact with the sand. Two swabs were collected per individual. One swab was placed in a sterile Eppendorf tube and stored at -20 °C within 10 hours of collection, referred to as the 'dry swab'. The second sample type, referred to as 'wet swab', was placed in a sterile Eppendorf tube containing 0.5 ml BuccalFix Buffer (Isohelix) and shaken. Theoretically, BuccalFix Buffer allows for stable storage at ambient temperatures. At the end of the nesting season, samples were transported back to the UK on ice and stored until extraction. After oviposition turtles were tagged with PIT (AVID) tags to avoid multiple sampling (Cameron *et al.*, 2019, Stiebens *et al.*, 2013b). Curved carapace length (CCL) and curved carapace width (CCW) were measured (±0.1 cm), and the presence of the leech parasite, *O. margoi*, on the cloaca was recorded (Lockley *et al.*, 2020). Additionally, two sand samples were collected from Sal and Santo Antão, samples were collected randomly within the nesting beach at ~ 30–35 cm, digging by hand using sterile gloves.
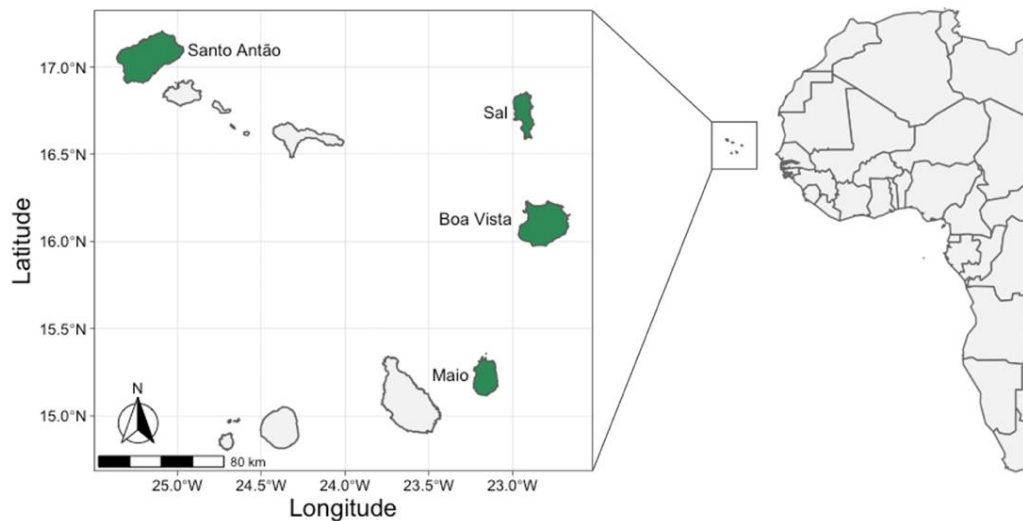
Figure 5.1: Map of the Cabo Verde archipelago in the eastern Atlantic, approximately 570 km from the coast of Senegal, West Africa. Islands where samples were collected are highlighted in green.

5.3.2 DNA extraction

DNA was extracted from swab samples using DNeasy 96 Blood & Tissue Kit (QIAGEN) with the following protocol modifications. Proteinase K, Buffer ATL, and Buffer AL–ethanol steps were carried out in 2 ml Eppendorf tubes. For dry swabs, we doubled the volume of proteinase K, Buffer ATL, and Buffer AL–ethanol (40 μl, 360 μl, 820 μl respectively), to cover the entire swab during lysis. This was not necessary for the wet swabs stored in BuccalFix Buffer. Samples were left to lyse overnight. For sand samples, 0.25 g of sand was aseptically transferred to a sterile 2 ml Eppendorf and extracted using the same modifications as dry swabs. Each extraction plate had 2 blanks, containing no swab sample: one to be used as an extraction blank throughout the extraction process, amplification and sequencing; the other as a PCR blank. For all samples, the elution step was conducted twice with 75 μl of warmed elution buffer (50 °C) incubated for 5 minutes each time before centrifuging, to increase DNA yield.

5.3.3 PCR amplification and sequencing

We amplified the V4 region of 16S ribosomal RNA gene (390 bp) using the primers   515F   (5'-GTGYCAGCMGCCGCGGTAA-3')   –   806R   (5'-

GGACTACNVGGGTWTCTAAT-3', (Apprill *et al.*, 2015, Parada *et al.*, 2016). PCR products were amplified using 5 µl 5x PCRBIO HiFi buffer, 1 µl of both primers (5 pmol/µl), 0.25 µl PCRBIO HiFi Polymerase (2u/µl), 16.75 µl PCR grade H_2O and 1 µl Template DNA for a total volume of 25 µl. Each reaction was conducted in duplicate on a Bio-Rad T100 Thermal Cycler using the following protocol: 1 minute at 95 °C, 15 seconds at 95 °C (x34), 15 seconds at 65 °C (x34), 30 seconds at 72 °C (x34), 5 minutes at 72 °C. Amplification was validated using gel electrophoresis. Duplicated reactions were pooled prior to sequencing. Extraction blanks were handled in the same way as samples, but PCR blanks had 1 µl Template DNA substituted with 1 µl PCR grade H_2O.

5.3.4 Illumina sequencing data analysis

Amplicons were sent to the Genome Centre, London, for PCR product clean up, Illumina MiSeq library prep and 300 bp paired-end sequencing. The quality of demultiplexed paired-end sequencing reads (R1 and R2) was evaluated using FastQC (Andrews, 2010). Primer and adapter sequences were removed using Trimmomatic (Bolger *et al.*, 2014). We trimmed sequences to minimise sequencing errors whilst maximising read retention, using the bioinformatics tool, *Figaro* v1.0.0, to estimate trimming thresholds (White *et al.*, 2008). *Figaro* parameters were set to an estimated amplicon size of 390 bp, and a minimum overlap of 70 bp, allowing up to 100 bp to be trimmed from the overlapping region, resulting in 30 bp and 78 bp being removed from the 3' end of the sequence for R1s and R2s respectively. QIIME 2 v2020.02 (Bolyen *et al.*, 2019) and DADA2 software packages (Callahan *et al.*, 2016) were used to analyse the pre-processed reads into amplicon sequence variants (ASVs) using default parameters. Taxonomy was assigned using Greengenes 13_8 database (McDonald *et al.*, 2012). ASVs have recently become the preferred method of characterising 16S rRNA gene sequences over operational taxonomic units (OTUs) as OTUs cluster sequences by

167

sequence similarity (typically at a 97% similarity threshold) which may mask true biological variation and can be troublesome to compare across studies (Edgar, 2017). ASVs are exact sequence variants that are comparable across studies improving reproducibility (Callahan *et al.*, 2017).

5.3.5 Statistical analyses

Statistical analyses were conducted in R version 3.6.3, packages used include: 'phyloseq', 'vegan', 'lmerTest', 'lme4' (R Core Team, 2020). Normality of model residuals were tested and transformed if required to meet test assumptions. All models were backward selected using Akaike's Information Criterion (AIC) values to retain the optimal reduced model.

ASVs classified as mitochondria or chloroplasts, and those that were present in only one sample, were removed from the dataset. Additionally, samples with fewer than 500 reads were removed from our dataset. Sand samples (N = 2) and blanks (N = 1 after filtering out samples with <500 reads) were not included in statistical analyses but were used for visual calibration and estimating risks of contamination. All analyses were conducted on non-rarefied data and data rarefied to 2000 reads in order to test for the consistency of the detected patterns. Rarefaction curves for both non-rarefied and rarefied datasets were plotted to verify saturation had been reached. To test for the correlation between turtle size and microbial community composition, a size category was assigned to each turtle in relation to their CCL. Turtles with a CCL in the lower tertile were classified as small, those in the upper tertile were classified as large, the remaining individuals were classified as medium.

5.3.5.1 Core microbiome

We examined the core microbiome of turtle cloaca samples to look for common ASVs across turtle nesting islands, storage methods and rarefied and non-rarefied

datasets. ASVs were classified as part of the core microbiome if they were present in at least 65% of all samples. Core ASVs were identified separately for the dry and the wet swabs in both rarefied and non-rarefied datasets.

5.3.5.2 Identifying the determinants of microbial diversity and structure.

Due to observed difference between storage methods and the paired nature of the swab samples, dry and wet swabs were split and analysed separately for our main biological questions. Firstly, we used LMMs to test for correlations between microbial diversity (Shannon, Gini-Simpson and PD) and turtle nesting island, the presence of the leech parasite (infected vs non-infected), CCL size groups, as well as their interactions. Extraction plate ID was set as a random factor. Tukey post-hoc tests between significant factors were used for pairwise comparisons.

To investigate the determinant of microbial community composition, we used PERMANOVAs based on Bray-Curtis, weighted UniFrac and unweighted UniFrac distance matrices. We compared microbial community composition between turtle nesting islands, the presence of the leech parasite, CCL size groups, and all their interactions. Pairwise PERMANOVAs were calculated between significant factors using the *pairwise.adonis2* function from the pairwiseAdonis package (Martinez Arbizu, 2020). All significant effects were followed by a dispersion test to ascertain the observed differences were not related to differences in data dispersion. Following the PERMANOVA, SIMPER and Kruskal-Wallis tests were conducted to identify which ASVs contribute most to variation among significant groups. Where significant relationships between microbial community and a given variable were identified, a non-metric multidimensional scaling (NMDS) plot was used to visualise the result.

5.3.5.3 Testing for storage methods

To test for differences among storage methods, a linear mixed effects model (LMM) was used to compare microbial diversity between swab types with extraction plate ID set as a random factor. Shannon's diversity index, Gini-Simpson index and Phylogenetic Diversity (PD) were used as measures of microbial diversity. Using PERMANOVAs based on Bray-Curtis dissimilarity matrix, weighted UniFrac and unweighted UniFrac distance matrices, we tested for differences in microbial composition between swab types. All significant effects were followed by a dispersion test with *betadisper* (vegan package) to ascertain that the observed differences were not related to differences in data dispersion. To identify which ASVs contributed most to the differences between swab types, we conducted similarity percentage analysis (SIMPER) using the *simper.pretty* function, to highlight ASVs that were associated with the between group variations (Steinberger, 2018). Simper tests were followed by Kruskal-Wallis tests using the function *kruskal.pretty* to test for their statistical significance (Steinberger, 2018).

To estimate whether one of the sampling methods yielded better sequence information, we used a LMM to compare non-rarefied sequence quality scores (Phred-33 scores) of each swab type, with extraction plate ID set as a random factor. Additionally, we used a LMM to compare the number of non-rarefied reads (log-transformed) between swab types, again with extraction plate ID set as a random factor.

5.4 Results

After filtering, 1,691,245 high-quality reads were retained from 234 samples, including 127 dry, and 104 wet samples (Table 5.1). 44 samples were removed as they had fewer than 500 reads (Supplementary Table 5.1). As expected, extraction blanks and all three PCR blanks contained extremely low read numbers, showing the reliability of

the protocols to avoid contamination. After rarefying to 2000 reads, 175 samples were retained (Table 5.1), both non-rarefied and rarefied datasets were used for core microbiome identification, bacterial diversity and community composition analysis. Rarefaction curves for both non-rarefied and rarefied datasets reached saturation plateaus (Supplementary Figure 5.1).

We identified 1690 ASVs across all samples. On average 57 ± 34(SD) ASVs were observed in swab samples, while sand samples harboured more ASVs with an average of 75 ± 9, and the remaining extraction blank contained 17 ASVs (Figure 5.2A). A similar pattern was seen at the phylum level (Figure 5.2B). Turtle samples harboured 30 different phyla, with Proteobacteria (59%), Bacteroidetes (13%), Lentisphaerae (6%), Fusobacteria (5%), Firmicutes (5%) being the most abundant, whilst still showing a large amount of inter-individual variation (Figure 5.3). We found 69.26% of turtles sampled were infected with *O. margoi*. For CCL size grouping, turtles shorter than 77 cm were classified as small, 77 cm – 79.5 cm were medium, and large turtles were longer than 79.5 cm

Table 5.1: Number of 16S rRNA samples: dry swabs, wet swabs, sand and extraction blanks retained for non-rarefied and rarefied microbial diversity and community composition.

		Island					
		Boa Vista	*Maio*	*Sal*	*Santo Antão*	*NA*	*Total*
Non-rarefied	*Dry Swab*	26	28	48	25	-	127
	Wet Swab	21	21	42	20	-	104
	Sand	-	-	1	1	-	2
	Extraction Blank	-	-	-	-	1	1
	Total	47	49	91	46	1	234
Rarefied	*Dry Swab*	23	21	44	23	-	111
	Wet Swab	14	11	24	12	-	61
	Sand	-	-	1	1	-	2
	Extraction Blank	-	-	-	-	1	1
	Total	38	33	68	35	1	175
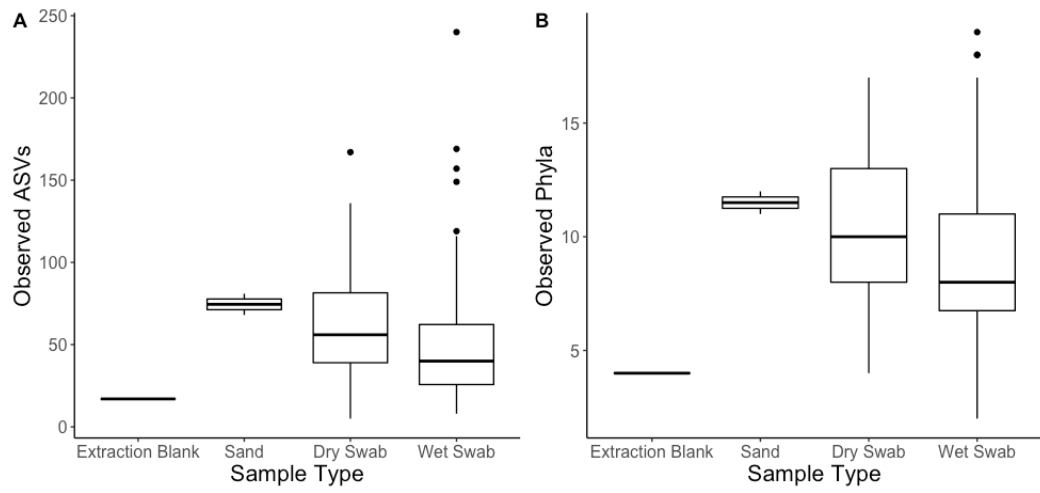
Figure 5.2: Variation in A) observed number of ASVs across sample types and B) observed number of microbial phyla among sample types. Non-rarefied data displayed.
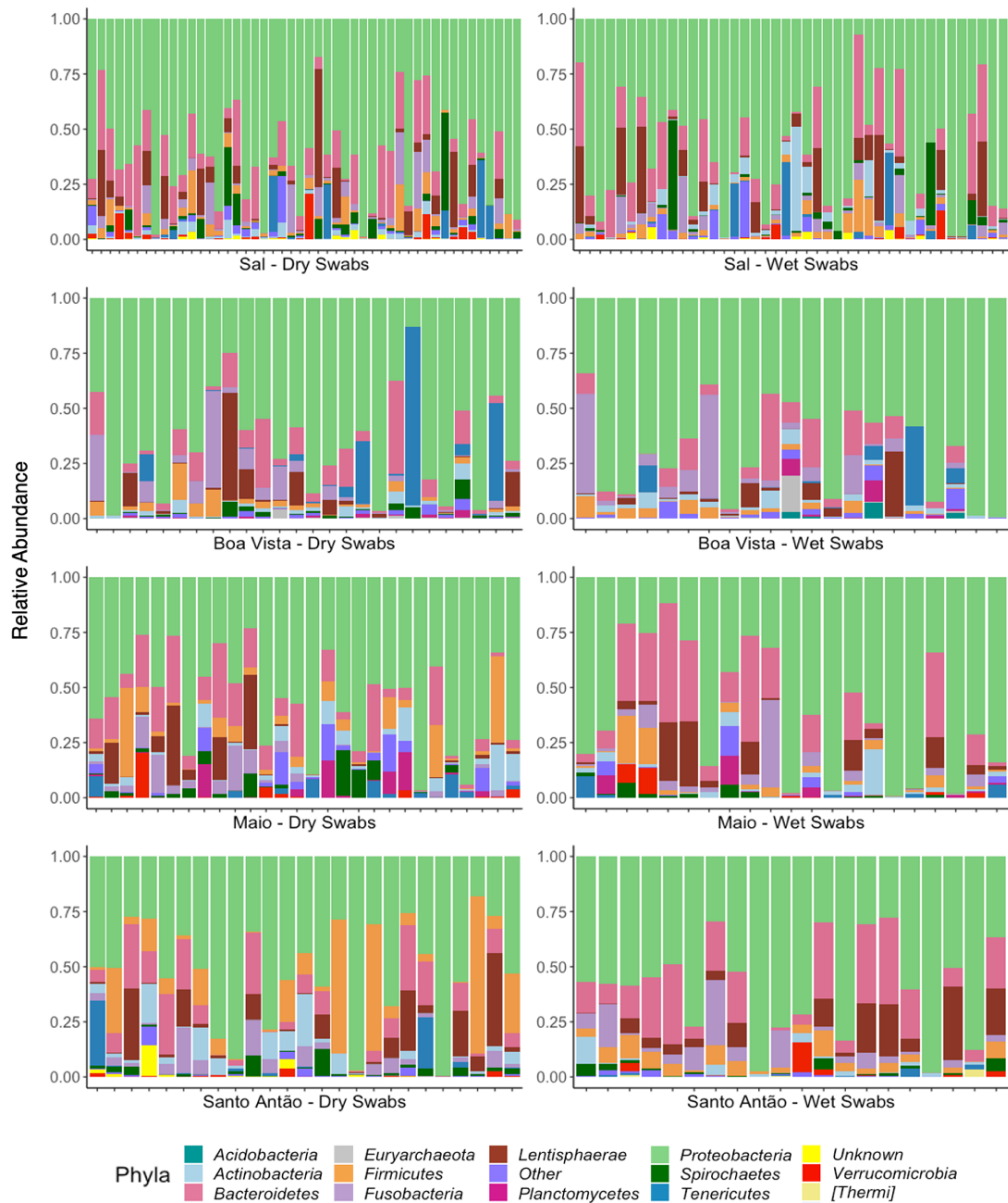
Figure 5.3: Relative abundance of microbial phyla identified in turtle samples categorised by island (Boa Vista, Maio and Santo Antão) and swab type (dry and wet). Bars represent individual turtle samples. The top 10 phyla observed in the individual are displayed, remaining phyla were grouped 'other'. Square brackets show the Greengenes database notation for proposed taxonomy.

5.4.1 Core microbiome

Seven ASVs were identified as core bacteria across turtles from all nesting islands at a 65% prevalence threshold (Table 5.2). Among these ASVs, three were shared among dry and wet swab samples: a *Helicobacter* species, an unclassified Oxalobacteraceae and an unclassified Vibrionales. Those three ASVs belonged to the phylum Proteobacteria. An additional three ASVs were detected solely in the dry samples (in both non-rarefied

and rarefied dataset). Those were *Propionibacterium acnes*, a *Brachyspira* species and *Photobacterium demsela*, belonging to Actinobacteria, Spirochaetes and Proteobacteria phyla, respectively. An unclassified Lentisphaerales ASV was identified in both dry swab sample datasets, and in the rarefied wet swab samples.

Table 5.2: Taxonomic information for seven core ASVs identified across non-rarefied and rarefied, dry and wet swabs. All taxa belong to the kingdom Bacteria. ASVs in bold were identified as core across all datasets analysed.

	Phylum	Class	Order	Family	Genus	Species
Non-rarefied Dry	**Proteobacteria**	**Epsilonproteobacteria**	**Campylobacterales**	**Helicobacteraceae**	***Helicobacter***	**-**
	Proteobacteria	**Betaproteobacteria**	**Burkholderiales**	**Oxalobacteraceae**	**-**	**-**
	Proteobacteria	**Gammaproteobacteria**	**Vibrionales**	**-**	**-**	**-**
	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	-	-	-
	Actinobacteria	Actinobacteria	Actinomycetales	Propionibacteriaceae	*Propionibacterium*	*acnes*
	Spirochaetes	[Brachyspirae]	[Brachyspirales]	Brachyspiraceae	*Brachyspira*	-
	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Photobacterium*	*damselae*
Non-rarefied Wet	**Proteobacteria**	**Epsilonproteobacteria**	**Campylobacterales**	**Helicobacteraceae**	***Helicobacter***	**-**
	Proteobacteria	**Betaproteobacteria**	**Burkholderiales**	**Oxalobacteraceae**	**-**	**-**
	Proteobacteria	**Gammaproteobacteria**	**Vibrionales**	**-**	**-**	**-**
Rarefied Dry	**Proteobacteria**	**Epsilonproteobacteria**	**Campylobacterales**	**Helicobacteraceae**	***Helicobacter***	**-**
	Proteobacteria	**Betaproteobacteria**	**Burkholderiales**	**Oxalobacteraceae**	**-**	**-**
	Proteobacteria	**Gammaproteobacteria**	**Vibrionales**	**-**	**-**	**-**
	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	**-**	**-**	**-**
	Actinobacteria	Actinobacteria	Actinomycetales	Propionibacteriaceae	*Propionibacterium*	*acnes*
	Spirochaetes	[Brachyspirae]	[Brachyspirales]	Brachyspiraceae	*Brachyspira*	-
	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Photobacterium*	*damselae*
Rarefied Wet	**Proteobacteria**	**Epsilonproteobacteria**	**Campylobacterales**	**Helicobacteraceae**	***Helicobacter***	**-**
	Proteobacteria	**Betaproteobacteria**	**Burkholderiales**	**Oxalobacteraceae**	**-**	**-**
	Proteobacteria	**Gammaproteobacteria**	**Vibrionales**	**-**	**-**	**-**
	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	**-**	**-**	**-**

5.4.2 Diversity and community structure inferred from dry swab samples

Given the philopatric nature of sea turtles and our objective to determine whether genetic and environmental factors are associated with microbial diversity and composition, we tested for correlations between microbial communities and turtle from different nesting islands, the presence of *O. margoi* and turtle size. For both non-rarefied and rarefied datasets, we found that microbial diversity of the dry swabs correlated with an interaction between turtle nesting island and *O. margoi* infection (LMM, non-rarefied: log-transformed Gini-Simpson: $F_{3,118} = 2.92$, $p < 0.05$; rarefied: Shannon: $F_{3,102} = 3.11$, $p < 0.05$ and log-transformed Gini-Simpson: $F_{3,102} = 2.82$, $p < 0.05$; Table 5.3, Figure 5.4). For the non-rarefied dataset, infected turtles nesting on Boa Vista had increased diversity in comparison to non-infected turtles, whilst turtles from Santo Antão showed the reverse patterns based on their infection status. No significant pairwise comparisons were identified when accounting for multiple testing (all post hoc Tukey tests $p > 0.05$, Supplementary Table 5.3). Phylogenetic diversity was not correlated with any of the tested variables regardless of rarefaction, and for the non-rarefied dataset, Shannon Diversity was also not related to any tested variables (Table 5.3). Nonetheless, overall, independently of the rarefied or the non-rarefied datasets, when significant correlations were found with microbial diversity, they involved nesting island and parasite infection (Table 5.3). Turtle curved carapace length was not correlated with microbial diversity.
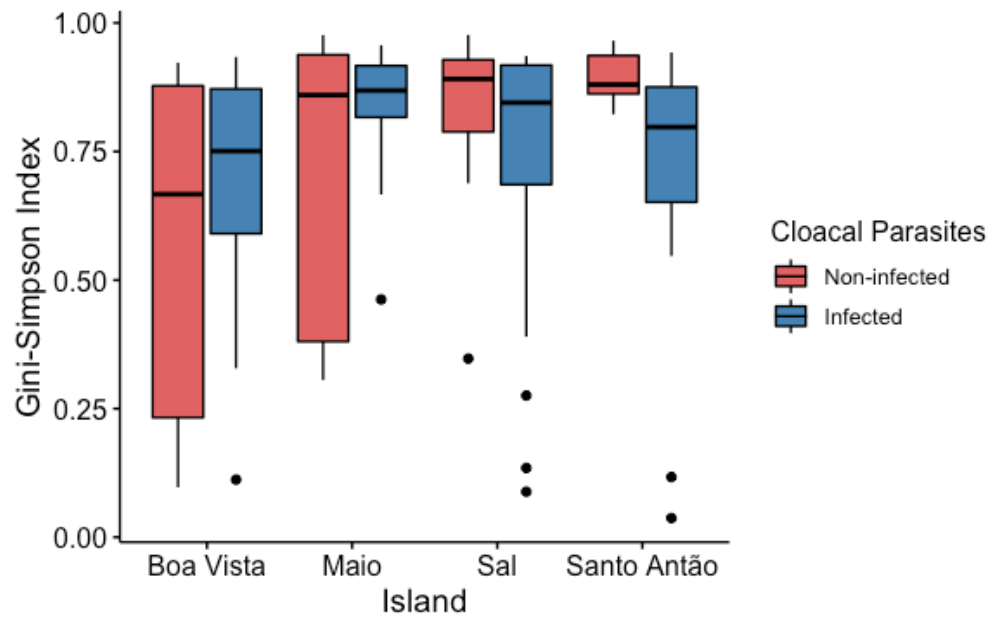
Figure 5.4: Microbial diversity of non-rarefied dry swab samples varied with an interaction between turtle nesting island and O. margoi infection (LMM, Gini-Simpson: $F_{3,118} = 2.92$, $p < 0.05$). No significant pairwise comparisons were identified.

Table 5.3: Summary table reporting best reduced models of the effect of turtle nesting island, O. margoi infection (Parasite) and CCL size group (CCL) along with their three-way interactions on Shannon diversity index, Gini-Simpson index and phylogenetic diversity. All models were backward selected using AIC. Some models ran on log-transformed data to meet the assumptions of the test. Significant results are in bold.

		Non-Rarefied				Rarefied		
		Df	F	p		Df	F	p
	Shannon Diversity Index				*Shannon Diversity Index*			
	Island	3,109	0.66	0.5758	Island	3,102	1.45	0.232
	Parasite	1,110	0.02	0.8848	Parasite	1,102	0.08	0.774
	CCL	1,110	0.10	0.747	**Island:Parasite**	**3,102**	**3.11**	**0.030**
	Island:Parasite	3,108	0.42	0.7359				
	Island:CCL	3,109	0.77	0.5156				
	Parasite:CCL	1,110	0.02	0.8786				
	Island:Parasite:CCL	3,108	0.48	0.6969				
Dry Swabs	*Gini-Simpson Index (log-transformed)*				*Gini-Simpson Index (log-transformed)*			
	Island	3,118	2.54	0.060	Island	3,102	1.32	0.273
	Parasite	1,118	0.00	0.996	Parasite	1,102	0.00	0.980
	Island:Parasite	**3,118**	**2.92**	**0.037**	**Island:Parasite**	**3,102**	**2.82**	**0.043**
	Phylogenetic Diversity				*Phylogenetic Diversity*			
	Island	3,109	0.15	0.932	Island	3,94	0.36	0.784
	Parasite	1,110	0.02	0.900	Parasite	1,94	0.06	0.813
	CCL	1,110	0.00	0.956	CCL	1,94	0.16	0.686
	Island:Parasite	3,107	0.44	0.727	Island:Parasite	3,94	0.19	0.902
	Island:CCL	3,109	0.17	0.916	Island:CCL	3,94	0.39	0.764
	Parasite:CCL	1,110	0.03	0.860	Parasite:CCL	1,94	0.08	0.774
	Island:Parasite:CCL	1,107	0.44	0.727	Island:Parasite:CCL	3,94	0.21	0.890
	Shannon Diversity Index				*Shannon Diversity Index*			
	Island	3,86	1.12	0.348	Island	3,43	1.37	0.264
	Parasite	1,85	1.98	0.164	Parasite	1,43	0.11	0.739
	CCL	1,85	0.47	0.494	CCL	1,43	0.04	0.851
	Island:Parasite	3,86	0.50	0.681	Island:Parasite	3,43	0.76	0.520
	Island:CCL	3,86	1.19	0.319	Island:CCL	3,43	1.39	0.259
	Parasite:CCL	1,85	1.90	0.172	Parasite:CCL	1,43	0.08	0.783
	Island:Parasite:CCL	3,86	0.52	0.668	Island:Parasite:CCL	3,43	0.75	0.529
Wet Swabs	*Gini-Simpson Index (log-transformed)*				*Gini-Simpson Index (log-transformed)*			
	Island	3,86	0.28	0.840	Island	3,43	0.48	0.697
	Parasite	1,85	0.88	0.351	Parasite	1,43	0.29	0.593
	CCL	1,85	0.07	0.792	CCL	1,43	0.01	0.920
	Island:Parasite	3,86	0.10	0.958	Island:Parasite	3,43	0.11	0.954
	Island:CCL	3,86	0.33	0.804	Island:CCL	3,43	0.48	0.699
	Parasite:CCL	1,85	0.85	0.360	Parasite:CCL	1,43	0.24	0.626
	Island:Parasite:CCL	3,86	0.12	0.951	Island:Parasite:CCL	1,43	0.09	0.965
	Phylogenetic Diversity				*Phylogenetic Diversity*			
	Island	3,86	0.30	0.826	Island	3,43	0.81	0.495
	Parasite	1,85	0.01	0.903	Parasite	1,43	0.02	0.899
	CCL	1,85	0.07	0.794	CCL	1,43	0.28	0.601
	Island:Parasite	3,86	0.57	0.638	Island:Parasite	1,43	1.34	0.274
	Island:CCL	3,86	0.30	0.824	Island:CCL	3,43	0.82	0.492
	Parasite:CCL	1,85	0.00	0.953	Parasite:CCL	1,43	0.01	0.941
	Island:Parasite:CCL	3,86	0.59	0.626	Island:Parasite:CCL	3,43	1.32	0.280

Turtles nesting on different islands tended to harbour different microbial communities. For the non-rarefied dataset, we found a significant effect of nesting island on microbial community composition, regardless of distance metric (PERMANOVA, Unweighted UniFrac: $F_{3,125} = 1.69$, $p < 0.001$, Weighted UniFrac: $F_{3,125} = 1.51$, $p < 0.05$ and Bray-Curtis: $F_{3,125} = 1.39$ $p < 0.05$, Figure 5.5A, Table 5.4). Pairwise PERMANOVAs showed that the microbial community of turtles from Santo Antão differed significantly from those of turtles from all other islands (pairwise PERMANOVA, Unweighted UniFrac: Santo Antão – Boa Vista: $F_{1,48} = 2.21$, $p < 0.01$, Santo Antão – Sal: $F_{1,71} = 2.38$, $p < 0.01$ and Santo Antão – Maio: $F_{1,51} = 2.06$, $p < 0.05$, Supplementary Table 5.4). Consistent with non-rarefied data, microbial community composition of dry swabs using the rarefied dataset was also significantly different among turtles nesting on the different islands (PERMANOVA, Unweighted UniFrac: $F_{3,109} = 1.45$, $p < 0.01$, Table 5.4); however, when rarefied, a significant relationship was found for the unweighted UniFrac metric only. For the rarefied dataset, community composition of Santo Antão turtles significantly differed from turtles nesting on Boa Vista and Sal (Santo Antão - Boa Vista: $F_{1,43} = 1.90$, $p < 0.05$, Santo Antão - Sal: $F_{1,56} = 2.04$, $p < 0.01$). Noteworthy, data dispersion may explain a part of the observed pattern for rarefied dry swabs, however data dispersion is homogenous for non-rarefied dry swabs (betadisper, Table 5.4).

We found eight ASVs for the non-rarefied dataset and seven ASVs for the rarefied dry swab dataset that significantly differed in abundance among turtles from different nesting islands. For the non-rarefied dataset, five of these ASVs belonged to the phylum Proteobacteria, the most dominant phylum in our samples (Table 5.6, Supplementary Table 5.5). Furthermore, we identified an ASV belonging to the Lentisphaeria class that was consistently in higher abundance in turtles from Santo Antão in comparison to turtles from other nesting islands (Supplementary Table 5.5). Focusing on Santo Antão turtles

further revealed two Proteobacteria ASVs differed in abundance with Boa Vista turtles, and two Proteobacteria ASVs with Maio turtles. Lastly, one Bacteroidetes ASV was more common in Santo Antão than in Sal and Maio turtles (Table 5.5, Supplementary Table 5.5). Other ASVs differed in abundance among turtles from different islands. For instance, a Proteobacteria and Tenericutes ASVs differ between turtles nesting on Boa Vista and Sal turtles (Table 5.5, Supplementary Table 5.5). The complete pairwise comparison is reported in Table Supplementary Table 5.5. In the rarefied dataset, five of the seven ASVs identified by SIMPER were the same as those identified in the non-rarefied dataset. These were the Pasteurellaceae family ASV, the Lentisphaerales order ASV, *Shewanella algae*, *Campylobacter* species and a Paraprevotellaceae family ASV (Table 5.6, Supplementary Table 5.6). The slight difference between ASVs highlighted by the SIMPER analyses of each dataset shows the potential impact of the storage method is rather minimal.

When investigating the link between microbiomes and parasite infection, we found that the presence of *O. margoi* was correlated with changes in host microbial community, with more overall inter-individual variation detectable in non-infected turtles. This was the case for both the non-rarefied and rarefied datasets, but only when using the unweighted UniFrac distance metric (PERMANOVA, non-rarefied: $F_{1,125}=$ 1.60, $p < 0.05$; rarefied: $F_{1,109} = 1.79$, $p < 0.05$; Figure 5.5B, Table 5.4), We found no differences in data dispersion between infected and non-infected groups (betadisper, all β-diversity metrics: $p > 0.05$, Table 5.4). Five ASVs were found in significantly different abundances between turtles infected with *O. margoi* and those that were not for the non-rarefied dataset from the SIMPER analysis (Table 5.6, Supplementary Table 5.7). Specifically, a *Fusobacterium* species, two ASVs belonging to the class Gammaproteobacteria and an ASV belonging to the family Bacteroidaceae were significantly more abundant in infected turtles, whilst one ASV, a *Helicobacter* species,

was more abundant in non-infected turtles. For rarefied data, the difference between infected and non-infected turtles was mostly associated with three ASVs (Table 5.6, Supplementary Table 5.7) of these, a *Helicobacter* species and an ASV assigned to the Cardiobacterials order were also identified in non-rarefied dry swab samples. Interestingly for the rarefied dataset from dry swabs, no significant effects of nesting island, turtle size, or *O. margoi* infection on microbial composition were detected when using weighted UniFrac or Bray-Curtis distances.
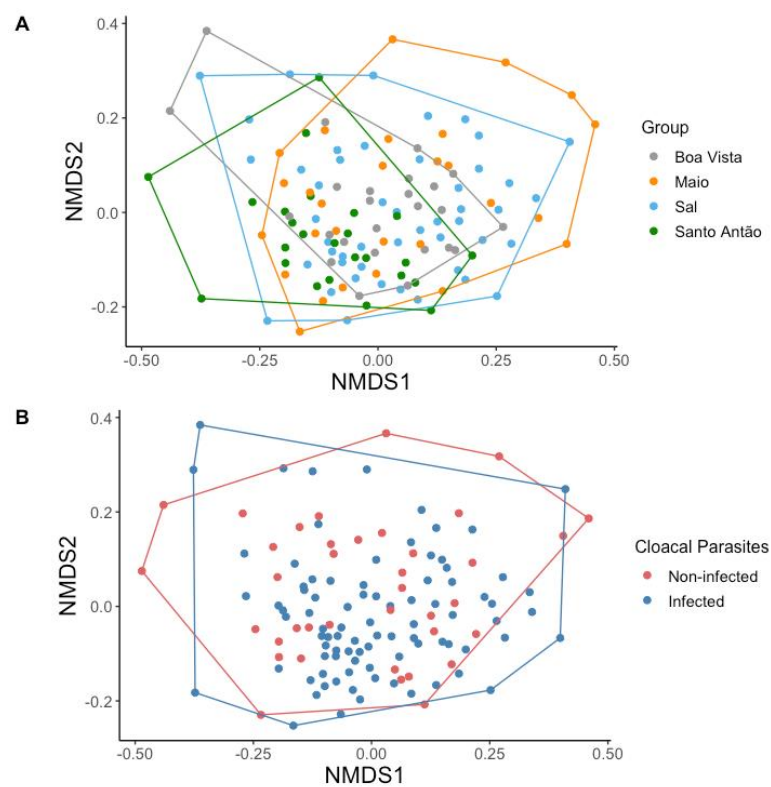
Figure 5.5: Non-metric multidimensional scaling (NMDS) plot of the microbial community within non-rarefied dry swab samples A) based on unweighted UniFrac by island (PERMANOVA, $F_{3,125} = 1.69$, $p < 0.001$) and B) the microbial community based on unweighted UniFrac by O. margoi infection (PERMANOVA, $F_{1,125} = 1.60$, $p < 0.05$).

Table 5.4: Summary of all PERMANOVA and betadisper results. Significant results are in bold.

Dataset	Factor	Df	Unweighted UniFrac Distance			Weighted UniFrac Distance			Bray-Curtis Dissimilarity		
			F-value	*R2*	*P-value*	*F-value*	*R2*	*P-value*	*F-value*	*R2*	*P-value*
Non-Rarefied Wet and Dry Swabs	**Swab Type**	**1**	**5.03**	**0.02**	**0.001**	**2.32**	**0.10**	**0.014**	**4.35**	**0.02**	**0.001**
	Residuals	229	-	0.98	-	-	0.99	-	-	0.98	-
	Total	230	-	1.00	-	-	1.00	-	-	1.00	-
	Dispersion Test	*Df*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*
	Swab Type dispersion	1,229	1.81	999	0.180	0.07	999	0.788	0.16	999	0.692
Rarefied Wet and Dry Swabs	**Swab Type**	**1**	**1.82**	**0.01**	**0.047**	**2.52**	**0.01**	**0.012**	**1.80**	**0.01**	**0.028**
	Residuals	170	-	0.99	-	-	0.99	-	-	0.99	-
	Total	171	-	1.00	-	-	1.00	-	-	1.00	-
	Dispersion Test	*Df*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*
	Swab Type dispersion	1,170	0.20	999	0.660	0.99	999	0.320	0.89	999	0.346
Non-Rarefied Dry Swabs	**Island**	**3**	**1.69**	**0.04**	**0.001**	**1.51**	**0.04**	**0.048**	**1.39**	**0.03**	**0.013**
	Cloacal Parasites	**1**	**1.59**	**0.01**	**0.036**	0.93	0.01	0.486	**1.69**	**0.01**	**0.023**
	CCL size group	2	0.91	0.01	0.650	0.58	0.01	0.941	0.99	0.02	0.463
	Island:Cloacal Parasites	3	0.94	0.02	0.618	1.16	0.03	0.264	1.13	0.03	0.186
	Island:CCL size group	6	0.99	0.05	0.518	0.99	0.05	0.470	1.11	0.05	0.170
	Cloacal Parasites:CCL size group	2	1.00	0.02	0.456	1.02	0.02	0.419	0.77	0.01	0.895
	Island:Cloacal Parasites:CCL size group	5	1.01	0.04	0.459	0.93	0.04	0.566	1.09	0.04	0.208
	Residuals	103	-	0.81	-	-	0.82	-	**-**	0.81	**-**
	Total	125	-	1.00	-	-	1.00	-	-	1.00	-
	Dispersion Test	*Df*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*
	Island dispersion	**3,122**	**3.47**	**999**	**0.018**	*1.83*	*999*	*0.144*	*1.41*	*999*	*0.242*
	Cloacal Parasite dispersion	1,124	3.36	999	0.071	-	-	-	*0.09*	*999*	*0.767*
Rarefied Dry Swabs	**Island**	3	**1.45**	**0.04**	**0.008**	1.36	0.04	0.116	1.28	0.03	0.084
	Cloacal Parasites	1	**1.66**	**0.01**	**0.043**	0.96	0.01	0.434	1.66	0.01	0.053
	CCL size group	2	0.84	0.02	0.789	0.65	0.01	0.871	1.13	0.02	0.292
	Island:Cloacal Parasites	3	1.14	0.03	0.184	1.16	0.03	0.236	1.20	0.03	0.163
	Island:CCL size group	6	1.04	0.06	0.321	1.03	0.06	0.390	1.17	0.06	0.151
	Cloacal Parasites:CCL size group	2	1.16	0.02	0.208	0.99	0.02	0.439	0.82	0.01	0.783

	Factor	Df	F-value	N.Perm	Pr(>F)	F-value	N.Perm	Pr(>F)	F-value	N.Perm	Pr(>F)
	Island:Cloacal Parasites:CCL size group	5	0.97	0.04	0.546	0.98	0.04	0.485	1.12	0.05	0.213
	Residuals	87	-	0.78	-	-	0.79	-	-	0.77	-
	Total	109	-	1.00	-	-	1.00	-	-	1.00	-
	Dispersion Test	*Df*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*
	Island dispersion	**3,106**	**3.34**	**999**	**0.022**	-	-	-	-	-	-
	Cloacal Parasite dispersion	1,108	0.24	999	0.627	-	-	-	-	-	-
Non-Rarefied Wet Swabs	Island	3	1.03	0.03	0.379	1.12	0.03	0.298	1.23	0.04	0.083
	Cloacal Parasites	**1**	**1.56**	**0.02**	**0.040**	1.35	0.01	0.205	**1.59**	**0.02**	**0.028**
	CCL size group	2	1.12	0.02	0.260	1.06	0.02	0.368	1.00	0.02	0.467
	Island:Cloacal Parasites	3	0.94	0.03	0.578	1.25	0.04	0.175	1.11	0.03	0.239
	Island:CCL size group	6	1.08	0.06	0.231	**1.63**	**0.09**	**0.008**	**1.24**	**0.07**	**0.020**
	Cloacal Parasites:CCL size group	2	0.87	0.02	0.704	0.68	0.01	0.827	0.85	0.02	0.787
	Island:Cloacal Parasites:CCL size group	5	0.92	0.05	0.716	0.77	0.04	0.892	0.94	0.05	0.659
	Residuals	80	-	0.78	-	-	0.76	-	-	0.77	-
	Total	102	-	1.00	-	-	1.00	-	-	1.00	-
	Dispersion Test	*Df*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*
	Cloacal Parasite dispersion	1,101	1.27	999	0.260	-	-	-	*0.10*	999	0.758
	Island:CCL size group dispersion	11,92	-	-	-	1.09	999	0.380	**2.59**	**999**	**0.007**
Rarefied Wet Swabs	**Island**	3	1.07	0.05	0.323	1.50	0.07	0.089	**1.43**	**0.07**	**0.039**
	Cloacal Parasites	1	**2.20**	**0.04**	**0.012**	**2.19**	**0.03**	**0.033**	**1.72**	**0.03**	**0.032**
	CCL size group	2	0.96	0.03	0.510	1.07	0.03	0.378	0.99	0.03	0.467
	Island:Cloacal Parasites	3	1.41	0.07	0.053	1.20	0.05	0.220	1.11	0.05	0.240
	Island:CCL size group	6	0.95	0.09	0.600	1.49	0.13	0.066	1.16	0.11	0.128
	Cloacal Parasites:CCL size group	2	1.26	0.04	0.149	1.22	0.04	0.253	1.02	0.03	0.412
	Island:Cloacal Parasites:CCL size group	5	1.22	0.10	0.103	1.31	0.10	0.110	1.05	0.08	0.354
	Residuals	37	-	0.59	-	-	0.55	-	-	0.59	-
	Total	59	-	1.00	-	-	1.00	-	-	1.00	-
	Dispersion Test	*Df*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*
	Island dispersion	3,56	-	-	-	-	-	-	**3.40**	**999**	**0.024**
	Cloacal Parasite dispersion	1,58	1.38	999	0.246	0.08	999	0.783	0.24	999	0.630

Table 5.5: Number of ASVs found in significantly different abundance (SIMPER) between turtles across different nesting islands.

	Boa Vista	Sal	Maio	Santo Antão
Boa Vista	-	2	4	3
Sal	-	-	2	2
Maio	-	-	-	4
Santo Antão	-	-	-	-

5.4.3 Diversity and community structure inferred from wet swab samples

Compared to the samples stored in dry conditions, those stored in BuccalFix Buffer yielded a lower overall average number of ASVs, but all 30 phyla were represented (Figure 5.2 and Figure 5.3). Unlike the dry swab samples, we found no significant correlation between microbial diversity of wet swab samples and the nesting island of a turtle or the presence of cloacal parasites for any of the diversity indices used (Table 5.3). Like the dry swab samples, microbial diversity of wet swab samples was unrelated to turtle size. These results for the wet swab samples were independent of rarefaction (Table 5.3).

Microbial community of non-rarefied wet swab samples correlated significantly with an interaction between the nesting island of a turtle and its size (PERMANOVA, Weighted UniFrac: $F_{1,102} = 1.62$, $p < 0.01$, and Bray-Curtis: $F_{1,102} = 1.24$ $p < 0.05$, Figure 5.6, Table 5.4). A pairwise PERMANOVA showed that small and large turtles from Maio carried significantly different microbial communities for both weighted UniFrac and Bray-Curtis distance metrics (pairwise PERMANOVA, Weighted UniFrac: $F_{1,13} = 1.83$, $p < 0.05$ and Bray-Curtis: $F_{1,13} = 1.53$ $p < 0.05$, Supplementary Table 5.8). Medium-sized turtles from Santo Antão harboured significantly different communities to small and large turtles from the same nesting island, but the bacterial community of small and large turtles were not significantly different from each other (pairwise PERMANOVA, Weighted UniFrac: Medium vs Small $F_{1,8} = 2.36$, $p < 0.05$, Medium vs Large $F_{1,15} = 2.72$. $p < 0.05$,

Small vs Large $F_{1,14}$ = 0.58. p = 0.824, Supplementary Table 5.8). We found no differences in data dispersion between turtle size groups (betadisper, all β-diversity metrics: p > 0.05, Table 5.4). In total, 29 ASVs showed differences in abundance across turtle nesting islands and size groups (Table 5.6, Supplementary Table 5.9). Interestingly, focusing on turtles nesting on Boa Vista, three ASVs showed to be different amongst turtles from different size groups, a *Fusobacterium*, a *Vibrio*, and an unclassified Actinobacteria (Supplementary Table 5.9). Turtles from Santo Antão had the most ASVs with significantly different abundances between turtles of different size, with five ASVs specifically explaining the community difference between large and medium turtles, all of which were in higher abundance in large turtles (Supplementary Table 5.9).

Additionally, the presence of *O. margoi* correlated with different microbial communities between infected and non-infected turtles, with the noticeable fact that non-infected turtles clustered more tightly compared with infected individuals (PERMANOVA, Unweighted UniFrac: $F_{1,102}$ = 1.56, p < 0.05, and Bray-Curtis: $F_{1,102}$ = 1.60, p < 0.05, Figure 5.7, Table 5.4). We found no differences in data dispersion between infected and non-infected individuals (betadisper, all β-diversity metrics: p > 0.05, Table 5.4). Specifically, we found that infected individuals had five ASVs in significantly higher abundance than non-infected turtles, *Paludibater* and unclassified Leptotrichiaceae, Carbiobacteriales, Bacteroidales and Lentisphaerales (Table 5.6, Supplementary Table 5.10).

Overall, rarefied wet swab samples revealed similar patterns to non-rarefied samples, with turtles' nesting island and parasite infections being the best variables to explain differences in microbial communities (PERMANOVA, Bray-Curtis: nesting island - $F_{3,59}$ = 1.43, p < 0.05 and cloacal parasites - $F_{1,59}$ = 1.67, p < 0.05, Table 5.4). For this rarefied dataset, there were 10 ASVs driving differences in microbial communities in turtles from different nesting islands (Table 5.6, Supplementary Table 5.11) and six

explaining the differences between infected and non-infected turtles. Four of those six were also identified in the non-rarefied dataset (an unclassified Cardiobacteriales, Lentisphaerales, Leptotrichiaceae and Paludibacter, Table 5.6, Supplementary Table 5.12).

Overall, our results show that independently of the dataset used, rarefied vs. non-rarefied, microbial community composition is consistently structured by turtle nesting island and the presence of the leech parasite, *O. margoi*. An effect of turtle size on microbial composition was only detected in the wet swab dataset, and this pattern was not consistent across all datasets.
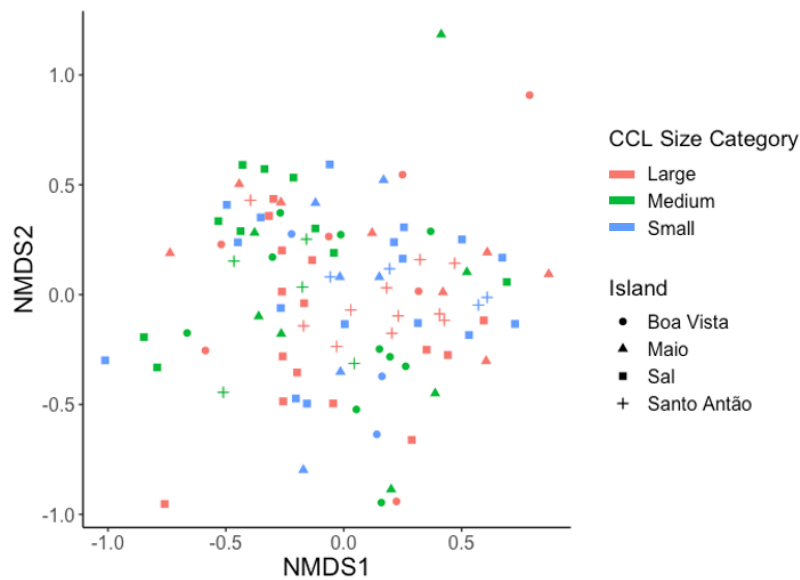
Figure 5.6: Non-metric multidimensional scaling (NMDS) plot of the microbial community within non-rarefied wet swab samples based on Bray-Curtis by island and CCL size category (PERMANOVA, $F_{1,102}$ = 1.26 p < 0.05).
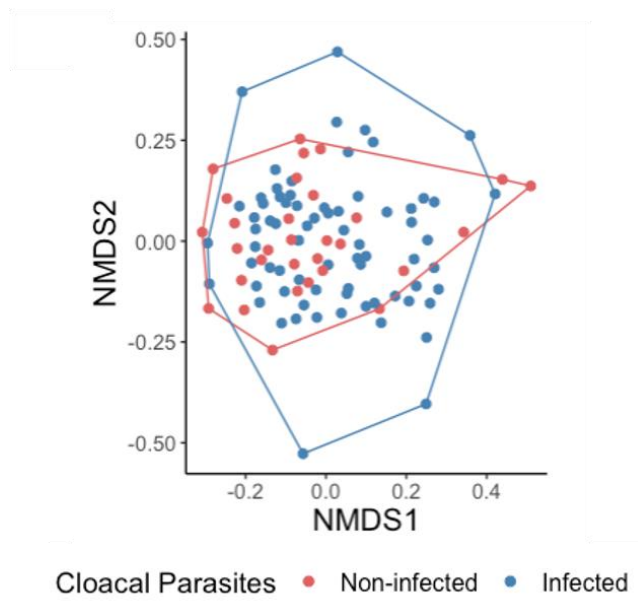
Figure 5.7: Non-metric multidimensional scaling (NMDS) plot of the microbial community within non-rarefied wet swab samples based on Bray-Curtis distance by O. margoi infection (PERMANOVA, $F_{1,102}$ = 1.60, p < 0.05).

5.4.4 Comparison among sample storage methods

Focusing on the non-rarefied dataset, we investigated whether the storage method, frozen or in BuccalFix Buffer, altered the microbial diversity. Firstly, we found wet swabs had significantly lower sequence quality scores (Phred-33 scores) than dry swabs (wet: 33.12 ± 0.7 (SD), dry: 33.45 ± 0.6, LMM, $F_{1,229}$= 11.00, p < 0.01, Figure 5.8A). Wet swab samples on average had fewer high-quality reads after filtering than dry swabs (LMM, $F_{1,229}$= 21.01, p < 0.001, wet: 6101 ± 7162 reads, dry: 8256 ± 5966 reads, Figure 5.8B).

Secondly, we found that phylogenetic diversity (PD) differed between our two collection methods, with dry swab samples having higher diversity (11.93 ± 4.28) than wet swab samples (9.62 ± 4.33, LMM, PD: $F_{1,229} = 20.4$, p < 0.001, Figure 5.9). No difference between swab type was detected when using Shannon diversity or Gini-Simpson index (LMM, Shannon: $F_{1,229} = 0.02$, p = 0.88, Gini-Simpson: $F_{1,229} = 0.06$, p = 0.80, Figure 5.9). After rarefying to 2000 reads, no differences in the microbial diversity between dry and wet swabs were identified (LMM, Shannon: $F_{1,169} = 0.71$, p = 0.4, Gini-Simpson: $F_{1,169} = 0.06$, p = 0.81, PD: $F_{1,169} = 3.06$, p = 0.08).
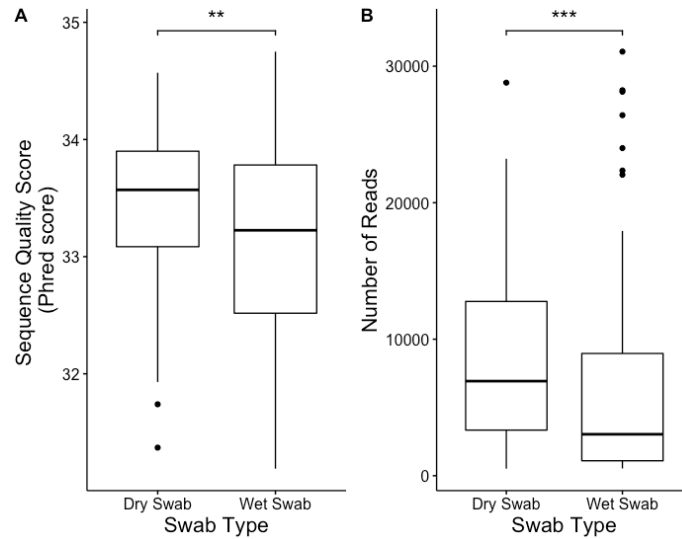
Figure 5.8: Boxplots showing A) higher average sequence quality scores in dry swab samples (LMM, $F_{1,229}$= 11.00, p < 0.01) and B) higher average number of reads for dry swab samples (LMM, $F_{1,229}$= 21.01, p < 0.001). Asterisks show significant comparisons.
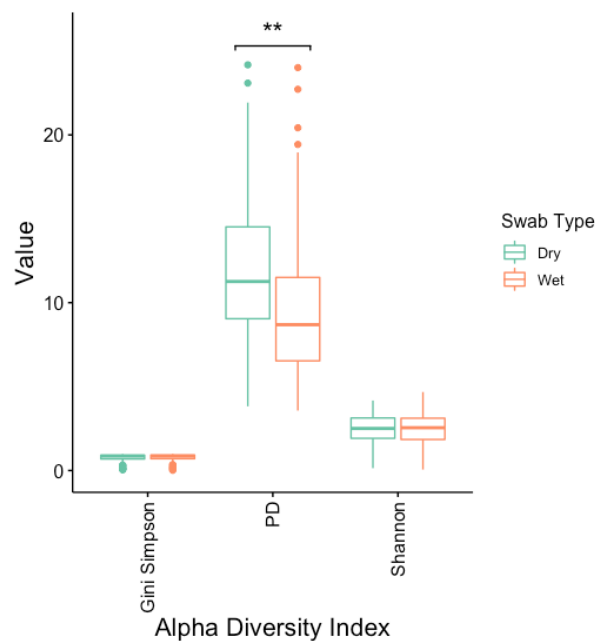
Figure 5.9: Sample storage method resulted in significantly higher phylogenetic diversity (PD) in dry swabs than wet swabs, other alpha diversity indexes were not different (LMM, PD: $F_{1,229}$ = 20.4, p < 0.001, Shannon: $F_{1,229}$ = 0.02, p = 0.88, Gini-Simpson: $F_{1,229}$ = 0.06, p = 0.80). Non-rarefied data displayed. Asterisks show significant comparisons.

Independently of the distance matrix used, we found that differences in microbial diversity translated into different bacterial community compositions between dry and wet swab samples on the non-rarefied dataset (PERMANOVA, Unweighted UniFrac: $F_{1,230}$ =

5.03, p < 0.01, Figure 5.10A, Weighted UniFrac: $F_{1,230} = 2.32$, p < 0.05, Figure 5.10B, Bray-Curtis: $F_{1,230} = 4.85$, p < 0.01, Figure 5.10C, Table 5.4), We found no differences in data dispersion between wet and dry swabs (betadisper, all β-diversity metrics: p > 0.05, Table 5.4). Using SIMPER analyses, we detected eight ASVs significantly contributing to the difference between non-rarefied wet and dry swabs (Table 5.6, see Supplementary Table 5.2 for mean abundance in each group and full taxonomic information). Of the ASVs identified, seven were found in greater abundance in dry swabs. These ASVs included *Photobacterium damselae,* one of the core microbes, two *Helicobacter* species, a *Vibrio* species, a *Brachyspira* species, an unclassified Bacteroidaceae and an unclassified Pasteurellaceae. The only ASV found in greater abundance in wet swab samples was an unclassified Rhizobiales.

We detected exactly the same results when focusing on the rarefied dataset. Microbial community composition was different between the storage methods independently of the distance matrix used (PERMANOVA, Unweighted UniFrac: $F_{1,171} = 1.82$, p < 0.05, Weighted UniFrac: $F_{1,171} = 2.52$, p < 0.05, Bray-Curtis: $F_{1,171} = 1.80$, p < 0.05, Table 5.4). We found no differences in data dispersion between wet and dry swabs (betadisper, all β-diversity metrics: p > 0.05, Table 5.4). A SIMPER analysis highlighted that three ASVs, a *Helicobacter* species, a *Brachyspira* species and an unclassified Rhizobiales were found in significantly different abundance between sample types (Table 5.6, Supplementary Table 5.2). These three ASVs overlap with those detected with the non-rarefied dataset, suggesting these ASVs strongly influence the difference in microbial community structure between wet and dry swab.

Due to the differences among storage methods that were identified, samples were analysed separately to understand the biological determinants of microbial diversity and community structure. Additionally, as each individual was swabbed twice the samples are paired by nature and if examined jointly, we would run the risk of pseudoreplication.
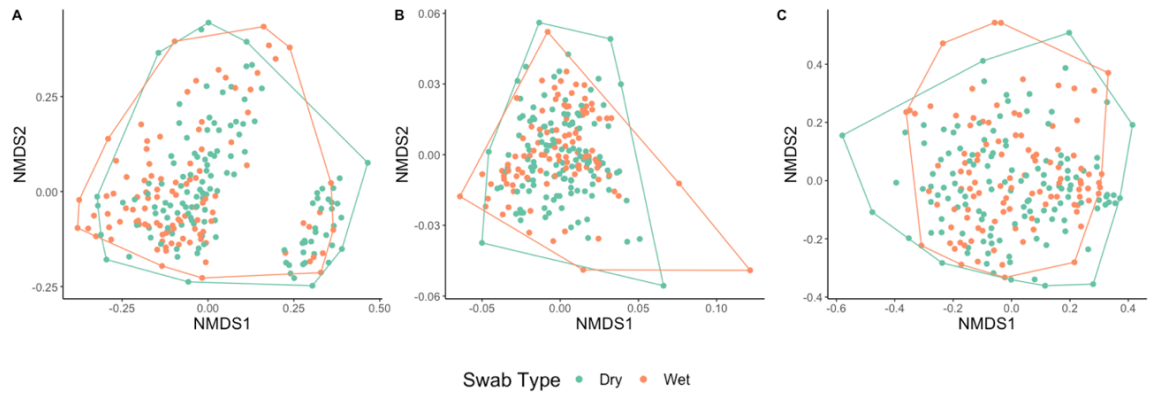
Figure 5.10: Non-metric multidimensional scaling (NMDS) plot of the microbial community of dry and wet cloacal swabs using A) Unweighted UniFrac distance (PERMANOVA, $F_{1,230} = 5.03$, $p < 0.01$), B) Weighted UniFrac distance (PERMANOVA, $F_{1,230} = 2.32$, $p < 0.05$), and C) Bray-Curtis dissimilarity (PERMANOVA, $F_{1,230} = 4.85$, $p < 0.01$).

Table 5.6: List of ASVs that are driving differences in microbial community at each analysis level, derived from SIMPERs carried out on variables that show significant variation in microbial community between groups. Ticks indicate that the abundance of the ASV differed significantly between at least one pair within SIMPER. Lowest taxonomic rank is the lowest classification obtained for ASV from Greengenes 13_8 database. Square brackets indicate the Greengenes database notation for proposed taxonomy.

| ASV | Lowest Taxonomic Rank | Dry vs Wet Swabs | | Dry Swabs | | | | | Wet Swabs | | |
| | | | | Island | | Parasites | | Island:CCL | Island | Parasites | |
		Non-rarefied data	Rarefied data	Non-rarefied data	Rarefied data	Non-rarefied data	Rarefied data	Non-rarefied data	Rarefied data	Non-rarefied data	Rarefied data
81554b0fce7cb99c919725c40ec6792a	[Paraprevotellaceae]	-	-	✔	✔	-	-	✔	✔	-	-
08b7ae36ca8ff80a868794ff4aa3af1c	Actinobacteria	-	-	-	-	-	-	✔	-	-	-
f0b70acd270f53b521eef73b3128a67f	*AF12* species	-	-	-	-	-	-	✔	-	-	-
95064b90639c6c20f85c141b5e8d6a15	Bacteroidaceae	✔	-	-	✔	✔	-	✔	-	-	-
21e65a24850031f756e25f0278d4d1af	Bacteroidales	-	-	-	-	-	-	✔	-	✔	-
6432d50fbe1bdecc92b8669a98cdcdcd	*Brachyspira* species	✔	✔	-	-	-	-	✔	-	-	-
413037c7bfae2c3535793bc67f4d9ae5	Campylobacter	-	-	✔	✔	-	-	✔	✔	-	-
76fd33f14910b1cd69086786956646b3	Cardiobacteriales	-	-	-	-	✔	✔	✔	✔	✔	✔
1aeef1011c39dc1e7b526ebeab794a68	Desulfovibrionaceae	-	-	-	-	-	-	-	✔	-	-
4d72007c70f4abc0bf9e8fe8b826d07d	Enterobacteriaceae	-	-	-	-	-	-	✔	-	-	-
5699b09d3c91c083e8ce6bcafc742e25	*Fusobacterium* species	-	-	-	-	✔	-	✔	-	-	✔
2cc1b1f5c5a88c4e0abc1b9f44a77ea4	Halobacteriaceae	-	-	-	-	-	-	✔	-	-	-
ad7f3efd185f47aab4f7badf2c822cc8	*Helicobacter* species	✔	-	-	-	-	-	✔	-	-	-
f62dc709536edba0135034e740915a92	*Helicobacter* species	✔	✔	✔	-	✔	✔	-	-	-	-
17fbebb3d9f3bbd3508cd5348ae7ea7c	Lentisphaerales	-	-	✔	✔	-	-	✔	✔	✔	✔
61beb22e53acd837370e43e68a012642	Lentisphaerales	-	-	-	-	-	-	✔	-	-	-
62312a4c1b699acbc6a6862c1a8bea83	Leptotrichiaceae	-	-	-	-	-	-	✔	-	✔	✔
b4bb370a59f2637d406ccc0f5a1b2a1e	*Morganella morganii*	-	-	-	-	-	-	-	-	-	✔

ID	Taxon	1	2	3	4	5	6	7	8	9	10
33c1fcdbc21a10a2a033f53d47db7893	*Mycoplasma* species	-	-	✔	-	-	-	✔	-	-	-
b501af89d0069bb6642c88a656b4d962	Neisseriaceae	-	-	-	-	-	-	✔	✔	-	-
dfe06fc469793226d7bd06f40fd3f4f5	*Odoribacter* species	-	-	-	-	-	-	✔	-	-	-
d1b122b873bef2a9246fbd7ead125ca7	*Ornithobacterium* species	-	-	-	-	-	✔	-	-	-	-
ce7504bf8df2af8e3139aacad6b79dae	Paludibacter	-	-	-	-	-	-	✔	-	✔	✔
ff0b033d6bfeaf9f7afadb421c67cfbc	*Parabacteroides* species	-	-	-	-	-	-	✔	-	-	-
0015124e5777b57174a1af7411703bbb	Pasteurellaceae	✔	-	✔	✔	-	-	✔	✔	-	-
d20b46e3c9d79a8e49a48f112fc03d4f	Peptostreptococcaceae	-	-	-	✔	-	-	-	-	-	-
67b0cf6af26b9fd959d860a52a65bc58	*Photobacterium damselae*	✔	-	-	-	-	-	✔	✔	-	-
429f2491c9a2f652c5561ef90c463a2d	*Plesiocystis* species	-	-	-	-	-	-	✔	-	-	-
5a7b179b1b45f0fe2282f260bf073f60	*Propionibacterium acnes*	-	-	-	-	-	-	✔	-	-	-
8c472893b0fcfde7a64a96606df16221	RFP12	-	-	-	-	-	-	✔	-	-	-
95c95bb6bdcc327d75c06f8b3472723f	Rhizobiales	✔	✔	-	-	-	-	-	-	-	-
34239f14f57eed8a2971a240df604a06	*Shewanella algae*	-	-	✔	✔	-	-	✔	-	-	-
9be79137deae3d1b144b96c452427950	*Vibrio* species	✔	-	-	-	-	-	✔	✔	-	-
0429a8a999c3238e12bbfaa1714d385e	Vibrionales	-	-	✔	-	✔	-	✔	✔	-	-

5.5 Discussion

Host bound microbes can aid their host through improved immune response or nutrient uptake, to increase condition and ultimately reproductive fitness (Bäckhead *et al.*, 2005, Gerardo and Parker, 2014, Kelly and Salinas, 2017, Turnbaugh *et al*., 2006). As such, they can alter life history traits and influence evolutionary trajectory (Brucker and Bordenstein, 2012, McFall-Ngai *et al.*, 2013, Rudman *et al.*, 2019). Here, we investigated the effects that turtle nesting island, parasite infection and turtle size had on the cloacal microbiome of wild loggerhead sea turtles. Generally, we identified Proteobacteria, Bacteroidetes, Lentisphaerae, Fusobacteria and Firmicutes as the dominant phyla in the cloacal microbiome, and found 7 ASVs that formed the core microbiome of turtles, regardless of the sampling location. More importantly, we found that both the nesting island of turtles and *O. margoi* infection, as well as their interaction, were the best descriptors of microbial diversity and community composition. Given the field nature of our study, we also tested for the differences in microbial diversity and community that could emerge from sample storage methods, suggesting dry samples to provide better sequencing results – though overall patterns remained qualitatively similar between the two methods. Our study shows that locally adapted nesting groups of sea turtles have differences in cloacal microbiome diversity and community, likely as the results of both genetic and environmental factors. This result is consistent with the evolution of local adaptation of the nesting groups, despite spending the vast majority of their life in a shared oceanic environment.

Regardless of the storage method and the datasets (rarefied vs non-rarefied), our overall results show broadly similar patterns and suggest turtles from different nesting groups harboured different microbial community composition. We find pairwise differences in the microbiomes between turtles from Santo Antão and those from all other nesting islands. Not only is Santo Antão the most geographically distant island, but turtles

nesting there are also the most genetically differentiated turtles from the core of the species distribution seen in Boa Vista, Sal and Maio islands (Baltazar-Soares *et al.*, 2020). Given that turtles are thought to not feed during the nesting period (Lutz *et al.*, 2002), and given that these turtles spend most of their lifetime in the same feeding ground, this result supports the idea that the cloacal microbiome of loggerhead sea turtles is, in part, influenced by host genetics. Differences in the cloacal microbiome of loggerhead sea turtles have been observed between rookies in the USA and Australia (Scheelings *et al.*, 2020), but no such differences have been described at an ecologically-relevant, island-specific scale. This is likely the result of the strong philopatric nature of turtles nesting in Cabo Verde, where distinct genetic groups have evolved broadly matching the nesting islands (Baltazar-Soares *et al.*, 2020, Stiebens *et al.*, 2013b). In Cabo Verde, nesting groups of turtles harbour island-specific diversity in the immune genes of the major histocompatibility complex (MHC) class I (Stiebens *et al.*, 2013b). The MHC, and particularly class I MHC genes, are known to affect the microbial composition of the gut microbiome in both humans and mice (Palma *et al.*, 2010, Vaahtovuo *et al.*, 2003), and this influence is likely to extend to other organisms such as the loggerhead sea turtle. We identified several ASVs that drive the differences in microbial communities among turtles from different nesting islands. For instance, we found a greater abundance of three ASVs, unclassified Lentisphaerales, Paraprevotellaceae, and Peptostreptococcaceae, in turtles nesting on Santo Antão, in comparison to turtles nesting on all other islands. Future studies would need to focus on the taxonomic resolution of these ASVs to test for their impact on the metabolism of sea turtles and identify whether different genetic basis, e.g. based on MHC class I genes, respond differently to colonization by those microbes. Another potential hypothesis may be the local environmental conditions (Spor *et al.*, 2011) could lead to microbial differences between the nesting island such as the sand-based microbiome which could be linked to the amount of human use which could

introduce novel bacterial species or colour of the sand leading to difference in temperatures.

Ecological and trophic niches are major factors of gut microbiome composition (Desai *et al.*, 2012, Faith *et al.*, 2011, Hildebrandt *et al.*, 2009, Parks *et al.*, 2013, Turnbaugh *et al.*, 2009, Wu *et al.*, 2011). It was recently shown that there are up to three different nesting strategies used by turtles nesting in Cabo Verde, with two oceanic strategies and a neritic one (Cameron *et al.*, 2019). For instance, turtles nesting on Santo Antão use the two oceanic strategies whereby they exploit both the habitat exposed to upwelling variations in the western coasts of Africa and a more characteristic oceanic habitat. On the other hand, on Boa Vista, in addition to the two oceanic strategies, turtles also exploit a neritic habitat on the coastline of Sierra Leone (Cameron *et al.*, 2019, Hawkes *et al*., 2006). If the feeding environment were the sole determinant of cloacal microbial diversity and composition, we would expect that sea turtles nesting in Boa Vista to be the most differentiated ones. Since this is not what we observed, we can parsimoniously exclude feeding ecology as a major determinant of the observed nesting group specific patterns. This reinforces the perspective that nesting-group specific microbiome stems from genetic effects associated with turtle local adaptation and philopatry.

Interestingly, we found that infection by *O. margoi* correlated with changes of microbial diversity and community composition. The microbial community of non-infected individuals clustered more tightly, which suggests more similar microbial communities. Greater microbial variation in infected turtles is probably because the impact parasites have on host microbiomes can differ greatly, with studies showing both increased (Lee *et al.*, 2014, Rosa *et al.*, 2018) and decreased microbial diversity (Houlden *et al.*, 2015, McKenney *et al.*, 2015). This may be correlated to direct interactions and manipulation of the microbiome by parasites (Dheilly *et al.*, 2015). Alternatively,

parasites also activate the host immune system and mucus production, impacting the microbiome indirectly (Leung *et al.*, 2018). Noteworthy, even the correlation between parasites and the cloacal microbiome was nesting group-specific. This matches our expectations because the prevalence of this parasite varies across Cabo Verde, with higher infection prevalence in turtles nesting in the east of the archipelago, at the core of the species distribution (Lockley *et al.*, 2020). Here as well, we can associate this change with nesting group-specific immunity, resulting in host-parasite-microbiota interactions demonstrating specific variation among nesting groups, and a potential role in the evolution of turtle local adaptation. The difference in microbiomes was associated with several ASVs, and the majority of these were more abundant in infected turtles. Parasites are able to modify an organism's microbiome by modulating the immune response of the host and changing the environment to best suit the requirements of the parasite (Dheilly *et al.*, 2015, Leung *et al.*, 2018, Shi *et al.*, 2014). It is plausible that the presence of *O. margoi* in the cloaca could alter the cloacal environment, resulting in cloacal dysbiosis and the colonization of possible pathogenic microbes, as we observed with the increased abundance of *Fusobacterium* species or *Vibrio* sp. Additionally leeches may introduce novel bacterial species which could explain the microbial variation between infected and non-infected turtles.

Turtles in our study had a high *O. margoi* prevalence (69.26%). It has been shown this parasite's prevalence in loggerhead sea turtles in Cabo Verde has increased from 10% in 2010 to 33% in 2017 (Lockley *et al.*, 2020). As rates of infection increase, it becomes increasingly important to understand the impact *O. margoi* has on the microbiome of turtles, to monitor for potential dysbiosis. While parasite infections can alter microbial communities; the microbiota can also protect the host against parasite infection (Holm *et al.*, 2015, Jaenike *et al.*, 2010, Oliveira-Sequeira *et al.*, 2014, White *et al.*, 2018). If the presence of certain parasites alters the microbiome, there may be additional knock-on

effects to the host health. Whilst the turtle, fibropapillomatosis causing, chelonid herpesvirus ChHV5 is yet to be found in loggerhead sea turtles in Cabo Verde, *O. margoi* is a potential vector (Greenblatt *et al.*, 2005, Jones *et al.*, 2016). Any reduction in host fitness due to this increased prevalence and altered microbiome could potentially increase the risk of ChHV5 virus infection.

While the patterns are less consistent across sample types and dataset, we identified a relationship between the nesting island of a turtle and turtle size with their microbial community composition. This size by nesting island interaction demonstrates both intrinsic and extrinsic determinants of microbial community composition. Both body mass and gut volume have been linked to increased microbial diversity previously (Godon *et al.*, 2016). Sea turtles continuously grow throughout their life (Omeyer *et al.*, 2017), therefore size can be used as a proxy for age. Age and development stages have been correlated with shifts in the host's diet, which would indirectly impact the microbiome (Ramirez *et al.*, 2015). Cameron *et al.* (2019) found that turtles nesting on the island of Boa Vista that use oceanic upwelling feeding strategy were larger than neritic turtles. Different feeding strategies could be a driver for the interaction we find between turtle nesting island and size. Exploring this interaction further by directly linking diet, through stable isotopes for instance, to the microbial diversity and community of loggerhead sea turtles would help to disentangle environment and the genetic effects.

If we are to learn about the microbiomes of wild populations, and particularly so of cryptic species like sea turtles which come to land only shortly for nesting, it is essential to identify relevant storage methods. Cryopreservation, freezing at -80 °C immediately after collection, has often been highlighted as the best approach to ensure microbiomes stay true to their original composition (Vandeputte et al., 2017). However, this is not always possible in remote locations, and storage at room temperature for an extended period may result in change in microbial diversity over time (Choo et al., 2015, Shaw et

al., 2016). As such, stabilisation buffers are often used to preserve samples, but these can sometimes reduce DNA quality and purity (Dominianni *et al.*, 2014, Vandeputte *et al.*, 2017). In this study, independently of the storage method, we identified phyla that have previously been identified in other populations of loggerhead turtles, e.g. Proteobacteria, Bacteroidetes, Lentisphaeare, Fusobacteria and Firmicutes (Abdelrhman *et al.*, 2016, Ahasan *et al.*, 2017, Foti *et al.*, 2009). This result suggests that both methods retained relevant information. Interesting, freezing outperformed the stabilisation buffer mostly on sequencing characteristics, i.e. sequence quality and read length. We also found a higher microbial diversity in dry swab samples compared to the wet ones, but this effect did not translate overall into identifying diversity differences between nesting groups and infection status. The differences were mostly at the community levels. If we had chosen to use only BuccalFix, we may have missed key patterns of the microbiome composition. For studies conducting first screens in the field, more controls would probably be required to confirm which method most accurately represented the true turtle microbiome, but it is evident that the stabilising buffer impacted sample quality and sequencing outcome.

Overall, our results indicate that microbial diversity and community of sea turtles differs among nesting group, their infection by a common parasite and even their size. Such host genotype-by-environment interactions describe well the patterns of local adaptation, whereby here, both turtles and their microbiome show signs of local adaptation as a result of philopatry. Identifying the genotype and feeding strategy of an individual could help to unravel the relative contribution of genetic or environmental factors on the cloacal microbiome. Yet, this study sets the baseline foundation to study the host-microbiome in a vulnerable and cryptic wild species.

Ethics statement

All animal sampling described were approved by the Ministério da Argicultra e Ambiente – Autroização number: 090/2019.

General Conclusions

Hosts and their microbiomes are in constant interaction and potentially coevolution, either through antagonist or mutualistic interactions (Herre *et al.*, 1999, Theis *et al.*, 2016, Van Valen, 1974), but the nature of this coevolution will vary depending on the strength of control imparted by the host on the microbes, the microbes on the host or the surrounding environmental conditions acting on them both. One outcome of these overall interactions could relate to host local adaptation (Kolodny and Schulenburg, 2020, Koskella and Bergelson, 2020, McFall-Ngai *et al.*, 2013, Rosenberg and Zilber-Rosenberg, 2016, Rudman *et al.*, 2019). Yet, what factors influence host-microbe interactions and local adaptation of wild vertebrate populations remain to be elucidated (Kohl *et al.*, 2018, Petipas *et al.*, 2020, Rennison *et al.*, 2019a, Sharon *et al.*, 2010).

The main objectives of this thesis were to increase our understanding of how microbial communities differ in response to both host genetics and environmental selection pressures and explore how these relationships may have coevolved in wild environments. Together with my collaborators, I examined how variation in host genetics (lineages and local genetic groups), parasite infection, and diet influenced the diversity and community structure of the microbiome of two locally-adapted wild host species: the three-spined stickleback, *Gasterosteus aculeatus* and the loggerhead sea turtle, *Caretta caretta*. Collectively, the findings of this thesis support the argument that coevolution can be detected in nature between a host and its microbiome, and that the structure of this coevolution is linked to a combination of host genetics and environmental pressures - parasites and diet in particular.

Our understanding of host-microbe interactions has increased rapidly in recent years (Bosch *et al.*, 2019, Gilbert *et al.*, 2012, McFall-Ngai *et al.*, 2013, Wu *et al.*, 2009), however, the majority of our knowledge comes from studies on captive animals or

laboratory-held model organisms (Amato, 2013, Hird, 2017). Indeed, only 14.3% of 650 publications between 2009 - 2016 on microbiomes had been carried out on wild organisms (Pascoe *et al.*, 2017). Laboratory studies are crucial, as they allow for the manipulation of a host and their microbes under controlled conditions and treatments, but changes in diet, increased stress exposure, and increased human contacts all have the potential to alter the structure of a host's microbiome (McKenzie *et al.*, 2017, Portz *et al.*, 2006, Uren Webster *et al.*, 2018). As a result, conclusions drawn from these studies may not be representative of host-microbe interactions in wild systems. This is because wild organisms are exposed to a vast array of selection pressures, including fluctuations in abiotic conditions, multiple parasite infections, social interactions and variation in food source availability - making host-microbiome interactions challenging to quantify in nature (Greyson-Gaito *et al.*, 2020, Hird, 2017). My thesis helps to fill this knowledge gap by exploring the wild microbiome of the three-spined stickleback (Chapters One & Three) and the loggerhead sea turtle (Chapter Four).

The reciprocal common garden experiment carried out in Chapters Three and Four is a key strength of this thesis. Whilst reciprocal common garden experiments are often logistically difficult due to the requirement of space, time, and resources, their results are highly informative. This approach allows for the study of genotype-by-environment effects, in the form of both the 'local vs. foreign' hypothesis as well as 'home vs. away' (de Villemereuil *et al.*, 2016, Hoban *et al.*, 2016, Kawecki and Ebert, 2004, Savolainen *et al.*, 2013). This design allowed us to disentangle the influence of genetics and the local environment that acts upon the host microbiome diversity and structure.

Firstly, we found evidence for local adaptation of host-microbe interactions through population-specific host microbiomes across geographical scales (Chapters Two and Five). In Chapter Two, we identified population-specific bacterial diversity and community structure across 11 three-spined stickleback populations in Europe and North

America. These results indicate a combined influence of a host's environment and genetics in structuring the microbiome. We conclude this, as if the fish microbiomes were purely influenced by the host environment, we would observe patterns of parallel evolution with ecotype-specific microbiomes, as the environmental pressures in two similar habitats, such as two lakes, would be more similar than two different habitats such as a lake and a river habitat. Therefore, population-specific patterns must be the result of localised environmental selection pressures and the increased genetic relatedness of individuals within a population. To frame these results in the context of previous research, population-specific microbiomes were also documented among stickleback populations from a single watershed in Canada, where the differences were associated with habitat type, habitat geomorphology and diet (Smith *et al.*, 2015). Our findings are an extension of previous work carried out on stickleback's microbiome as they extend across a larger geographical scale, including fish from North America and Europe. it is noteworthy that our results do not confirm all past patterns, as another study on the microbiome of the three-spined stickleback in Canada found evidence for parallel evolution between benthic and limnetic ecotype pairs (Rennison *et al.*, 2019a). This difference likely stems from the different selection pressures associated with the evolution of benthic-limnetic ecotype pairs compared to lake-river ecotypes.

We also identified nesting group-specific microbiomes across turtles nesting on different islands of the Cabo Verde archipelago (Chapter Five). To my knowledge, only one other paper has sampled the microbiome of sea turtles across geographically distinct nesting populations, with their findings showing differences in the microbial communities of loggerhead sea turtles originating from Florida, USA and Queensland, Australia (Scheelings *et al.*, 2020). My findings complement these and extend our understanding as the two populations sampled by Scheelings *et al.* (2020) are too genetically different to test for local adaptation, they also lack gene flow and shared feeding grounds, while

the populations within my study are sampled at an ecologically-relevant scale. The fact that I found evidence of local adaptation of the microbiome across two very distinct taxa suggests that this is likely to be a common evolutionary outcome of host-microbe interactions and should be explored in a wider range of organisms for further confirmation.

A finding of great interest from my research was the occurrence of host-parasite-microbe interactions, observed for both study species: in the three-spined stickleback (Chapter Four) and the loggerhead sea turtle (Chapter Five). Whilst both organisms are locally-adapted to a habitat type or nesting island (Baltazar-Soares *et al.*, 2020, Cameron *et al.*, 2019, DeFaveri and Merila, 2014, Hendry *et al.*, 2002, Lockley *et al.*, 2020, Stiebens *et al.*, 2013b), they otherwise have very different life histories. The consistent correlation between parasites and the microbiome across species shows that this is a crucial link to explore. The reciprocal nature of the interaction between parasites and microbes is only recently being better understood. A host's microbiome has the ability to defend its host against infection (Ford and King, 2016, King and Bonsall, 2017, King *et al.*, 2016), but some parasite-associated microbes also increase the ability of a parasite to infect its host (Adams *et al.*, 2006, Boemare and Akhurst, 2006). Additionally, parasites can interact directly with the host microbiome (Dheilly *et al.*, 2015), and can also activate the host immune system and mucus production, impacting the microbiome indirectly (Leung *et al.*, 2018). I have demonstrated that the interaction between a host, its parasites, and its microbiome can be observed when looking at (i) a single parasite species, *Ozobranchus margoi* infections within loggerhead sea turtles (Chapter Five) and (ii) at the parasite community level, composed of numerous individuals from various taxa (Chapter Four). The complexity of the latter study is not achievable under laboratory conditions, which further demonstrates how studies in wild systems are invaluable for understanding how other biological interactions influence host microbe-interactions.

It has been well established that host diet can influence the diversity and composition of their microbiome (Foster *et al.*, 2017, Li *et al.*, 2017, Muegge *et al.*, 2011, Sullam *et al.*, 2012). I confirmed this finding within the three-spined stickleback in Chapter Four, as clear host-microbe-diet interactions were detected. The influence of diet on the stickleback gut microbiome using stable isotopes has been identified previously, showing there is a large amount of inter-individual variation in the diet which correlated to changes within the gut microbiome (Bolnick *et al.*, 2014b, Bolnick *et al.*, 2014c). Due to the vast amount of research supporting the impact of diet on the microbiome, it is important to further explore the causality behind this correlation.

Chapter Three elucidated an important relationship that can be overlooked in host microbiome research: that feeding ecology is not independent of host-parasite resistance. The ecological interactions between parasite infection and feeding ecology are difficult to replicate *in vitro,* due to the complexity of both parasite and prey diversities, making our wild reciprocal common garden experiment an ideal way to explore this relationship. We found that a host's local environment, lineage, parasite community, and feeding ecology are correlated, showing the genotype-by-environment interactions that are a classic prerequisite of local adaptation (Kawecki and Ebert, 2004). This relationship between parasites and feeding ecology could be the result of different feeding strategies leading to variation in parasite exposure or, alternatively, could be a result of the infection whereby the host changes feeding ecology as a response to being infected (Brunner *et al.*, 2017, Locke *et al.*, 2014, Stutz *et al.*, 2014). Our use of stable isotopes allowed for long term patterns of feeding ecology to be evaluated (Lorrain *et al.*, 2002, Post, 2002), whilst previous research on the relationship between diet and host-parasite driven local adaptation has focused on gut content analysis (Bolnick *et al.*, 2020, Cirtwill *et al.*, 2016, Emde *et al.*, 2014, Kleinertz *et al.*, 2012, Reimchen and Nosil, 2001).

A difficulty I was confronted with over the course of my PhD research, is the rapid technical advancement in the field of host-microbe interactions. Whilst overall this can only improve the science, the speed at which the field moves results in a lack of standardised procedures for researching host-microbe interactions, rendering the comparison of results between studies challenging. Chosen storage methods, DNA extraction kits, 16S rRNA hypervariable regions, bioinformatic methods, and statistical analyses can all lead to differences in the patterns found from microbiome data, and it can be difficult to reliably compare across research papers (Arnold *et al.*, 2016, Clooney *et al.*, 2016, Kim *et al.*, 2017a, Laukens *et al.*, 2016, Lopez-Garcia *et al.*, 2018, Schloss, 2018). A key example of this lack of standardisation is the debate surrounding the rarefaction of data - some research groups suggest the removal of valid data is unadvisable, and methods to standardise data that do not require the removal of reads should be used (McMurdie and Holmes, 2014). In contrast, others argue that these methods do not result in a uniform number of reads, which can bias data analyses (McKnight *et al.*, 2019) and may not be sufficiently robust (Mandal *et al.*, 2015). When analysing Chapter Two data, at the start of my PhD, the vast majority of the literature rarefied their dataset to standardise the number of reads per sample, however, by the time I began work on my other chapters, the literature was more mixed and there was no general consensus. In order to overcome this potential limitation, we carried out analysis of both non-rarefied and rarefied data for Chapters Four and Five and confirmed that patterns identified were mostly consistent across both methods. Another technical discussion which animated intense debate relates to clustering sequences into Operational Taxonomic Units (OTUs). I used this common approach at the start of my PhD, however, by the end of it, there had been a paradigm shift towards not rarefying and assigning Amplicon Sequence Variations (ASVs). Whilst this change has improved the

reproducibility and comparability of microbiome studies, it does result in difficulties when comparing new and older findings (Callahan *et al.*, 2017, Edgar, 2017).

Were I to re-run my experiments again, an improvement I would consider would be to include positive controls in the microbiome analysis. While all my studies had negative controls, as was recommended by the literature, the use of positive controls, such as mock communities, which are mixtures of cultured organisms of known quantities, would allow for the verification of the accuracy of my chosen extraction and amplification methods (Salter *et al.*, 2014). Several groups have set out to standardise the field, such as the Earth Microbiome Project, the Human Microbiome Project, the Parasite Microbiome Project but inconsistencies still remain (Dheilly *et al.*, 2019, Gevers *et al.*, 2012, Gilbert *et al.*, 2014, Gilbert *et al.*, 2018, Schloss *et al.*, 2011, Turnbaugh *et al.*, 2007).

The exploration of the wild microbiome is crucial for improving our understanding of host-microbe interactions (Hird, 2017), however, studying wild organisms comes with its own set of limitations and disadvantages in comparison to controlled laboratory studies. Whilst the complexity of natural systems is a key factor of interest for host-microbe studies, it also results in a range of undocumented variables which may be of particular importance. Additionally, studies on the wild microbiome can only be correlative, inferring potential relationships but not causation (Bik, 2016). Additionally in laboratory-based studies there is the ability to track changes to a host's microbiome over time and in response to treatments. For example, even in our field experiments (Chapter Four and Five), we find the presence of parasites is correlated to changes in the microbiome, which are of great interest, however, we cannot infer which of these is driving the changes - does the parasite infection result in changes in the microbiome? Or do these differences in the microbiome allow for parasites to infect a host more easily? These questions can only be truly explored through controlled

laboratory manipulations and have been recently reviewed in (Stevens et al., 2021). Additionally, a similar set of questions can be asked concerning our findings from Chapter Three, as we found that parasite load, individual parasite taxa, and diet were correlated and likely coevolved. Nonetheless, it remains to be determined whether the changes in diet expose a host to different parasites leading to shifts in infections (Johnson *et al.*, 2009, Locke *et al.*, 2014, Stutz *et al.*, 2014), or whether the presence of those parasites drive a change in feeding ecology (Barber *et al.*, 2008, Lefevre *et al.*, 2009, Milinski, 1984, Ponton *et al.*, 2011). I believe that this shows that in order for the field of host-microbe interaction research to progress further, there is the need for a theoretical cross-talk to happen between laboratory and field studies.

As the number of descriptive studies on wild microbiomes grow, it becomes increasingly necessary to improve our understanding of the functions the microbiome plays within a range of hosts (Adair and Douglas, 2017). Hammer *et al.* (2019) suggested that hosts have a 'continuum of reliance' on symbionts, with some hosts requiring specific bacterial species to survive while others actively attempt to rid themselves of bacteria altogether. Questions for many species remain: are bacteria that are present in a host's microbiome purely commensal or transitioning through the gut with little impact on the host? Do they have a negative impact on the host? Or do they benefit the host through specific functions like pathogen defence and improved nutrient uptake? Computational approaches currently exist, such as PICRUSt, to predict the functionality of specific bacteria from 16S rRNA sequences, but how well they quantify functionality in novel environments is unknown (Langille *et al.*, 2013). Future research should explore the functionality of commonly occurring bacteria in wild vertebrate microbiomes, identifying which genes are expressed and metabolic processes are carried out by the microbiome as a whole as well as specific bacteria, using shotgun metagenomic approaches such as sequence-based or functional genomics (Sangwan *et al.*, 2016, Sommer *et al.*, 2009).

In summary, the studies conducted for my thesis have built on previous research and produced novel findings that enhance our scientific knowledge of wild host-microbe interactions. These findings highlight the importance of considering the role of the host microbiome in a host's evolution, as well as considering the complex interactions among a host, its environment, and its associated microbes and parasites. My hope is that this thesis will not only help improve our knowledge in the field of host-microbe interactions, but also our appreciation for the wonders of the microbial world and how with every new question the scientific community asks, their role and influence should be considered.

References

Aagaard, K. M. & Segars, J. H. 2014. What is the microbiome and how do we study it? *Semin Reprod Med,* 32**,** 3-4, 10.1055/s-0033-1361816.

Abdelrhman, K. F., Bacci, G., Mancusi, C., Mengoni, A., Serena, F. & Ugolini, A. 2016. A First Insight into the Gut Microbiota of the Sea Turtle Caretta caretta. *Front Microbiol,* 7**,** 1060, 10.3389/fmicb.2016.01060.

Adair, K. L. & Douglas, A. E. 2017. Making a microbiome: the many determinants of host-associated microbial community composition. *Curr Opin Microbiol,* 35**,** 23-29, 10.1016/j.mib.2016.11.002.

Adams, B. J., Fodor, A., Koppenhöfer, H. S., Stackebrandt, E., Patricia Stock, S. & Klein, M. G. 2006. Biodiversity and systematics of nematode–bacterium entomopathogens. *Biological Control,* 38**,** 4-21, 10.1016/s1049-9644(06)00126-5.

Ahasan, M. S., Waltzek, T. B., Huerlimann, R. & Ariel, E. 2017. Fecal bacterial communities of wild-captured and stranded green turtles (Chelonia mydas) on the Great Barrier Reef. *FEMS Microbiol Ecol,* 9310.1093/femsec/fix139.

Ahasan, M. S., Waltzek, T. B., Huerlimann, R. & Ariel, E. 2018. Comparative analysis of gut bacterial communities of green turtles (Chelonia mydas) pre-hospitalization and post-rehabilitation by high-throughput sequencing of bacterial 16S rRNA gene. *Microbiological Research,* 207**,** 91-99, 10.1016/j.micres.2017.11.010.

Aires, T., Serebryakova, A., Viard, F., Serrao, E. A. & Engelen, A. H. 2018. Acidification increases abundances of Vibrionales and Planctomycetia associated to a seaweed-grazer system: potential consequences for disease and prey digestion efficiency. *PeerJ,* 6**,** e4377, 10.7717/peerj.4377.

Al-Harbi, A. H. & Naim Uddin, M. 2004. Seasonal variation in the intestinal bacterial flora of hybrid tilapia (Oreochromis niloticus×Oreochromis aureus) cultured in earthen ponds in Saudi Arabia. *Aquaculture,* 229**,** 37-44, 10.1016/s0044-8486(03)00388-0.

Alberdi, A., Aizpurua, O., Bohmann, K., Zepeda-Mendoza, M. L. & Gilbert, M. T. P. 2016. Do Vertebrate Gut Metagenomes Confer Rapid Ecological Adaptation? *Trends in Ecology & Evolution,* 31**,** 689-699, 10.1016/j.tree.2016.06.008.

Altizer, S., Becker, D. J., Epstein, J. H., Forbes, K. M., Gillespie, T. R., Hall, R. J., Hawley, D. M., Hernandez, S. M., Martin, L. B., Plowright, R. K., Satterfield, D. A. & Streicker, D. G. 2018. Food for contagion: synthesis and future directions for studying host-parasite responses to resource shifts in anthropogenic environments. *Philosophical Transactions of the Royal Society of London B Biological Sciences,* 37310.1098/rstb.2017.0102.

Amann, R. I., Ludwig, W. & Schleifer, K. H. 1995. Phylogenetic Identification and In Situ Detection of Individual Microbial Cells without Cultivation. *Microbiological Reviews,* 59**,** 143-169.

Amato, K. R. 2013. Co-evolution in context: The importance of studying gut microbiomes in wild animals. *Microbiome Science and Medicine,* 110.2478/micsm-2013-0002.

Amir, A., Mcdonald, D., Navas-Molina, J. A., Kopylova, E., Morton, J. T., Zech Xu, Z., Kightley, E. P., Thompson, L. R., Hyde, E. R., Gonzalez, A. & Knight, R. 2017. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems,* 210.1128/mSystems.00191-16.

Anaya-Rojas, J. M., Brunner, F. S., Sommer, N., Seehausen, O., Eizaguirre, C. & Matthews, B. 2016. The association of feeding behaviour with the resistance and tolerance to parasites in recently diverged sticklebacks. *Journal of Evolutionary Biology,* 29**,** 2157-2167, 10.1111/jeb.12934.

Anders, S. & Huber, W. 2010. Differential expression analysis for sequence count data. *Nature Precedings,* 10.1038/npre.2010.4282.1.

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. Cambridge, UK: Babraham Institute.

Apprill, A. 2017. Marine Animal Microbiomes: Toward Understanding Host–Microbiome Interactions in a Changing Ocean. *Frontiers in Marine Science,* 410.3389/fmars.2017.00222.

Apprill, A., Mcnally, S., Parsons, R. & Weber, L. 2015. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology,* 75**,** 129-137, 10.3354/ame01753.

Arnold, J. W., Roach, J. & Azcarate-Peril, M. A. 2016. Emerging Technologies for Gut Microbiome Research. *Trends Microbiol,* 24**,** 887-901, 10.1016/j.tim.2016.06.008.

Arrington, D. A., Winemiller, K. O. L., W. F. & Akin, S. 2002. How often do fishes "run on empty"? *Ecology,* 8310.2307/3072046.

Austin, B. 2006. The bacterial microflora of fish, revised. *ScientificWorldJournal,* 6**,** 931-45, 10.1100/tsw.2006.181.

Avershina, E., Lundgard, K., Sekelja, M., Dotterud, C., Storro, O., Oien, T., Johnsen, R. & Rudi, K. 2016. Transition from infant- to adult-like gut microbiota. *Environmental Microbiology,* 18**,** 2226-36, 10.1111/1462-2920.13248.

Bäckhead, F., Ley, R., Sonnenburg, J. L., Peterson, D. A. & Gordon, J. I. 2005. Host-Bacterial Mutualism in the Human Intestine. *Science,* 307.

Bakke, T. A., Harris, P. D., Jansen, P. A. & Hansen, L. P. 1992. Host specificity and dispersal strategy in gyrodactylid monogeneans, with particular reference to Gyrodactylus saiaris (Platyhelminthes, Monogenea). *Diseases of Aquatic Organisms,* 13**,** 63-74, 10.3354/dao013063.

Baldo, L., Riera, J. L., Tooming-Klunderud, A., Alba, M. M. & Salzburger, W. 2015. Gut Microbiota Dynamics during Dietary Shift in Eastern African Cichlid Fishes. *PLoS ONE,* 10**,** e0127462, 10.1371/journal.pone.0127462.

Baltazar-Soares, M., Klein, J. D., Correia, S. M., Reischig, T., Taxonera, A., Roque, S. M., Dos Passos, L., Durao, J., Lomba, J. P., Dinis, H., Cameron, S. J. K., Stiebens, V. A. & Eizaguirre, C. 2020. Distribution of genetic diversity reveals colonization patterns and philopatry of the loggerhead sea turtles across geographic scales. *Sci Rep,* 10**,** 18001, 10.1038/s41598-020-74141-6.

Bar-On, Y. M., Phillips, R. & Milo, R. 2018. The biomass distribution on Earth. *Proceedings of the National Academy of Sciences of the United States of America,* 115**,** 6506-6511, 10.1073/pnas.1711842115.

Barber, I. 2007. Host-parasite interations of the three-spined sickleback. . *In:* Östlund-Nilsson, S., Mayer, I. & Huntingford, F. A. (eds.) *Biology of the Three-spined Stickleback.* London: CRC Press,.

Barber, I. 2013. Sticklebacks as model hosts in ecological and evolutionary parasitology. *Trends in Parasitology,* 29**,** 556-66, 10.1016/j.pt.2013.09.004.

Barber, I. & Scharsack, J. P. 2010. The three-spined stickleback - Schistocephalus solidus system: an experimental model for investigating host-parasite interactions in fish. *Parasitology,* 13710.1017/S0031182009991466.

Barber, I., Wright, H. A., Arnott, S. A. & Wootton, R. J. 2008. Growth and Energetics in the Stickleback Schistocephalus Host Parasite System: A Review of Experimental Infection Studies. *Behaviour,* 14510.1163/156853908792451412.

Barko, P. C., Mcmichael, M. A., Swanson, K. S. & Williams, D. A. 2018. The Gastrointestinal Microbiome: A Review. *J Vet Intern Med,* 32**,** 9-25, 10.1111/jvim.14875.

Barton, L. L. & Northup, D. E. 2011. Microbe–Microbe Interactions. *In:* Barton, L. L. & Northup, D. E. (eds.) *Microbial Ecology.*

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**,** 371-82, 10.1016/j.chom.2007.10.010.

Bell, M. A. & Foster, S. A. 1994. *The evolutionary biology of the threespine stickleback,* Oxford, Oxford University Press.

Berleman, J. E. & Kirby, J. R. 2009. Deciphering the hunting strategy of a bacterial wolfpack. *FEMS Microbiol Rev,* 33**,** 942-57, 10.1111/j.1574-6976.2009.00185.x.

Berner, D., Adams, D. C., Grandchamp, A. C. & Hendry, A. P. 2008. Natural selection drives patterns of lake-stream divergence in stickleback foraging morphology. *Journal of Evolutionary Biology,* 21**,** 1653-65, 10.1111/j.1420-9101.2008.01583.x.

Berner, D., Grandchamp, A. C. & Hendry, A. P. 2009. Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. *Evolution,* 63**,** 1740-53, 10.1111/j.1558-5646.2009.00665.x.

Berner, D., Roesti, M., Hendry, A. P. & Salzburger, W. 2010. Constraints on speciation suggested by comparing lake-stream stickleback divergence across two continents. *Molecular Ecology,* 19**,** 4963-78, 10.1111/j.1365-294X.2010.04858.x.

Berryman, A. A. 1992. The origins and evolution of predator-prey theory. *Ecology,* 73**,** 1530 - 1535.

Bertrand, M., Marcogliese, D. J. & Magnan, P. 2008. Trophic polymorphism in brook charr revealed by diet, parasites and morphometrics. *Journal of Fish Biology,* 72**,** 555-572, 10.1111/j.1095-8649.2007.01720.x.

Bestion, E., Jacob, S., Zinger, L., Di Gesu, L., Richard, M., White, J. & Cote, J. 2017. Climate warming reduces gut microbiota diversity in a vertebrate ectotherm. *Nat Ecol Evol,* 1**,** 161, 10.1038/s41559-017-0161.

Biagi, E., D'amico, F., Soverini, M., Angelini, V., Barone, M., Turroni, S., Rampelli, S., Pari, S., Brigidi, P. & Candela, M. 2019. Faecal bacterial communities from Mediterranean loggerhead sea turtles (Caretta caretta). *Environ Microbiol Rep,* 11**,** 361-371, 10.1111/1758-2229.12683.

Bik, E. M. 2016. The Hoops, Hopes, and Hypes of Human Microbiome Research. *Yale Journal of Biology and Medicine,* 89.

Bloodgood, J. C. G., Hernandez, S. M., Isaiah, A., Suchodolski, J. S., Hoopes, L. A., Thompson, P. M., Waltzek, T. B. & Norton, T. M. 2020. The effect of diet on the gastrointestinal microbiome of juvenile rehabilitating green turtles (Chelonia mydas). *PLoS ONE,* 15**,** e0227060, 10.1371/journal.pone.0227060.

Boemare, N. & Akhurst, P. 2006. The genera Photorhabdus and Xenorhabdus. *In:* Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & Stackebrandt, E. (eds.) *The Prokaryotes.* New York: Springer Science

Bolger, A. M., Lohse, M. & Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics,* 30**,** 2114-20, 10.1093/bioinformatics/btu170.

Bolnick, D. I., Barrett, R. D. H., Oke, K. B., Rennison, D. J. & Stuart, Y. E. 2018. (Non)Parallel Evolution. *Annual Reviews,* 49**,** 303-330, 10.1146/annurev-ecolsys-110617-062240.

Bolnick, D. I., Resetarits, E. J., Ballare, K., Stuart, Y. E. & Stutz, W. E. 2020. Scale-dependent effects of host patch traits on species composition in a stickleback parasite metacommunity. *Ecology***,** e03181, 10.1002/ecy.3181.

Bolnick, D. I., Snowberg, L. K., Caporaso, J. G., Lauber, C., Knight, R. & Stutz, W. E. 2014a. Major Histocompatibility Complex class IIb polymorphism influences gut microbiota composition and diversity. *Molecular Ecology,* 23**,** 4831-45, 10.1111/mec.12846.

Bolnick, D. I., Snowberg, L. K., Hirsch, P. E., Lauber, C. L., Knight, R., Caporaso, J. G. & Svanback, R. 2014b. Individuals' diet diversity influences gut microbial diversity in two freshwater fish (threespine stickleback and Eurasian perch). *Ecology Letters,* 17**,** 979-87, 10.1111/ele.12301.

212

Bolnick, D. I., Snowberg, L. K., Hirsch, P. E., Lauber, C. L., Org, E., Parks, B., Lusis, A. J., Knight, R., Caporaso, J. G. & Svanback, R. 2014c. Individual diet has sex-dependent effects on vertebrate gut microbiota. *Nat Commun,* 5**,** 4500, 10.1038/ncomms5500.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodriguez, A. M., Chase, J., Cope, E. K., Da Silva, R., Diener, C., Dorrestein, P. C., Douglas, G. M., Durall, D. M., Duvallet, C., Edwardson, C. F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J. M., Gibbons, S. M., Gibson, D. L., Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G. A., Janssen, S., Jarmusch, A. K., Jiang, L., Kaehler, B. D., Kang, K. B., Keefe, C. R., Keim, P., Kelley, S. T., Knights, D., Koester, I., Kosciolek, T., Kreps, J., Langille, M. G. I., Lee, J., Ley, R., Liu, Y. X., Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B. D., Mcdonald, D., Mciver, L. J., Melnik, A. V., Metcalf, J. L., Morgan, S. C., Morton, J. T., Naimey, A. T., Navas-Molina, J. A., Nothias, L. F., Orchanian, S. B., Pearson, T., Peoples, S. L., Petras, D., Preuss, M. L., Pruesse, E., Rasmussen, L. B., Rivers, A., Robeson, M. S., 2nd, Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S. J., Spear, J. R., Swafford, A. D., Thompson, L. R., Torres, P. J., Trinh, P., Tripathi, A., Turnbaugh, P. J., Ul-Hasan, S., Van Der Hooft, J. J. J., Vargas, F., Vazquez-Baeza, Y., Vogtmann, E., Von Hippel, M., Walters, W., et al. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol,* 37**,** 852-857, 10.1038/s41587-019-0209-9.

Bonder, M. J., Kurilshikov, A., Tigchelaar, E. F., Mujagic, Z., Imhann, F., Vila, A. V., Deelen, P., Vatanen, T., Schirmer, M., Smeekens, S. P., Zhernakova, D. V., Jankipersadsing, S. A., Jaeger, M., Oosting, M., Cenit, M. C., Masclee, A. A., Swertz, M. A., Li, Y., Kumar, V., Joosten, L., Harmsen, H., Weersma, R. K., Franke, L., Hofker, M. H., Xavier, R. J., Jonkers, D., Netea, M. G., Wijmenga, C., Fu, J. & Zhernakova, A. 2016. The effect of host genetics on the gut microbiome. *Nat Genet,* 48**,** 1407-1412, 10.1038/ng.3663.

Bordenstein, S. R. & Theis, K. R. 2015. Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes. *PLoS Biology,* 13**,** e1002226, 10.1371/journal.pbio.1002226.

Bosch, T. C. G., Guillemin, K. & Mcfall-Ngai, M. 2019. Evolutionary "Experiments" in Symbiosis: The Study of Model Animals Provides Insights into the Mechanisms Underlying the Diversity of Host-Microbe Interactions. *Bioessays,* 41**,** e1800256, 10.1002/bies.201800256.

Bouskra, D., Brezillon, C., Berard, M., Werts, C., Varona, R., Boneca, I. G. & Eberl, G. 2008. Lymphoid tissue genesis induced by commensals through NOD1 regulates intestinal homeostasis. *Nature,* 456**,** 507-510, 10.1038/nature07450.

Bowen, B. W., Bass, A. L., Chow, S. M., Bostrom, M., Bjorndal, K. A., Bolten, A. B., Okuyama, T., Bolker, B. M., Epperly, S., Lacasella, E., Shaver, D., Dodd, M., Hopkins-Murphy, S. R., Musick, J. A., Swingle, M., Rankin-Baransky, K., Teas, W., Witzell, W. N. & Dutton, P. H. 2004. Natal homing in juvenile loggerhead turtles (Caretta caretta). *Molecular Ecology,* 13**,** 3797-808, 10.1111/j.1365-294X.2004.02356.x.

Braendle, C., Miura, T., Bickel, R., Shingleton, A. W., Kambhampati, S. & Stern, D. L. 2003. Developmental origin and evolution of bacteriocytes in the aphid-Buchnera symbiosis. *PLoS Biology,* 1**,** E21, 10.1371/journal.pbio.0000021.

Brice, K. L., Trivedi, P., Jeffries, T. C., Blyton, M. D. J., Mitchell, C., Singh, B. K. & Moore, B. D. 2019. The Koala (Phascolarctos cinereus) faecal microbiome differs with diet in a wild population. *PeerJ,* 7**,** e6534, 10.7717/peerj.6534.

Britton, J. R. & Andreou, D. 2016. Parasitism as a Driver of Trophic Niche Specialisation. *Trends in Parasitology,* 32**,** 437-445, 10.1016/j.pt.2016.02.007.

Britton, J. R., Pegg, J. & Williams, C. F. 2011. Pathological and ecological host consequences of infection by an introduced fish parasite. *PLoS ONE,* 6**,** e26365, 10.1371/journal.pone.0026365.

Brucker, R. M. & Bordenstein, S. R. 2012. Speciation by symbiosis. *Trends in Ecology & Evolution,* 27**,** 443-51, 10.1016/j.tree.2012.03.011.

Brunner, F. S., Anaya-Rojas, J. M., Matthews, B. & Eizaguirre, C. 2017. Experimental evidence that parasites drive eco-evolutionary feedbacks. *Proceedings of the National Academy of Sciences of the United States of America,* 114**,** 3678-3683, 10.1073/pnas.1619147114.

Buckling, A. & Rainey, P. B. 2002. The role of parasites in sympatric and allopatric host diversification. *Nature,* 420**,** 496 - 499, doi.org/10.1038/nature01164.

Buffie, C. G., Bucci, V., Stein, R. R., Mckenney, P. T., Ling, L., Gobourne, A., No, D., Liu, H., Kinnebrew, M., Viale, A., Littmann, E., Van Den Brink, M. R., Jenq, R. R., Taur, Y., Sander, C., Cross, J. R., Toussaint, N. C., Xavier, J. B. & Pamer, E. G. 2015. Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. *Nature,* 517**,** 205-8, 10.1038/nature13828.

Burnham, K. P. & Anderson, D. R. 2004. Model selection: understanding AIC and multimodel inference, with contrasts to BIC. *ociological Methods and Research,,* 33**,** 261-304, 10.1177/0049124104268644.

Callahan, B. J., Mcmurdie, P. J. & Holmes, S. P. 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J,* 11**,** 2639-2643, 10.1038/ismej.2017.119.

Callahan, B. J., Mcmurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. & Holmes, S. P. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods,* 13**,** 581-3, 10.1038/nmeth.3869.

Callieri, C. 2008. Picophytoplankton in Freshwater Ecosystems: The Importance of Small-Sized Phototrophs. *Freshwater Reviews,* 1**,** 1-28, 10.1608/frj-1.1.1.

Cameron, S. J. K., Baltazar-Soares, M., Stiebens, V. A., Reischig, T., Correia, S. M., Harrod, C. & Eizaguirre, C. 2019. Diversity of feeding strategies in loggerhead sea turtles from the Cape Verde archipelago. *Marine Biology,* 16610.1007/s00227-019-3571-8.

Cano, J. M., Matsuba, C., Makinen, H. & Merila, J. 2006. The utility of QTL-Linked markers to detect selective sweeps in natural populations--a case study of the EDA gene and a linked marker in threespine stickleback. *Molecular Ecology,* 15**,** 4613-21, 10.1111/j.1365-294X.2006.03099.x.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Pena, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., Mcdonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J. & Knight, R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods,* 7**,** 335-6, 10.1038/nmeth.f.303.

Casale, P. & Tucker, A. D. 2017. *Caretta caretta* (amended version of 2015 assessment). *The IUCN Red List of Threatened Species,* 10.2305/IUCN.UK.2017-2.RLTS.T3897A119333622.en.

Chain, F. J., Feulner, P. G., Panchal, M., Eizaguirre, C., Samonte, I. E., Kalbe, M., Lenz, T. L., Stoll, M., Bornberg-Bauer, E., Milinski, M. & Reusch, T. B. 2014. Extensive copy-number variation of young genes across stickleback populations. *PLoS Genetics,* 10**,** e1004830, 10.1371/journal.pgen.1004830.

Chamberlain, S. A., Bronstein, J. L. & Rudgers, J. A. 2014. How context dependent are species interactions? *Ecology Letters,* 17**,** 881-90, 10.1111/ele.12279.

Chen, C., Huang, X., Fang, S., Yang, H., He, M., Zhao, Y. & Huang, L. 2018. Contribution of Host Genetics to the Variation of Microbial Composition of Cecum Lumen and Feces in Pigs. *Front Microbiol,* 9**,** 2626, 10.3389/fmicb.2018.02626.

Chen, S. W., Liu, C. H. & Hu, S. Y. 2019. Dietary administration of probiotic Paenibacillus ehimensis NPUST1 with bacteriocin-like activity improves growth performance and immunity against Aeromonas hydrophila and Streptococcus iniae in Nile tilapia (Oreochromis niloticus). *Fish & Shellfish Immunology,* 84**,** 695-703, 10.1016/j.fsi.2018.10.059.

Chen, W., Zhang, C. K., Cheng, Y., Zhang, S. & Zhao, H. 2013. A comparison of methods for clustering 16S rRNA sequences into OTUs. *PLoS ONE,* 8**,** e70837, 10.1371/journal.pone.0070837.

Chevalier, C., Stojanovic, O., Colin, D. J., Suarez-Zamorano, N., Tarallo, V., Veyrat-Durebex, C., Rigo, D., Fabbiano, S., Stevanovic, A., Hagemann, S., Montet, X., Seimbille, Y., Zamboni, N., Hapfelmeier, S. & Trajkovski, M. 2015. Gut Microbiota Orchestrates Energy Homeostasis during Cold. *Cell,* 163**,** 1360-74, 10.1016/j.cell.2015.11.004.

Chiu, K., Warner, G., Nowak, R. A., Flaws, J. A. & Mei, W. 2020. The Impact of Environmental Chemicals on the Gut Microbiome. *Toxicol Sci,* 176**,** 253-284, 10.1093/toxsci/kfaa065.

Choo, J. M., Leong, L. E. & Rogers, G. B. 2015. Sample storage conditions significantly influence faecal microbiome profiles. *Sci Rep,* 5**,** 16350, 10.1038/srep16350.

Cirtwill, A. R., Stouffer, D. B., Poulin, R. & Lagrue, C. 2016. Are parasite richness and abundance linked to prey species richness and individual feeding preferences in fish hosts? *Parasitology,* 143**,** 75-86, 10.1017/S003118201500150X.

Clayton, J. B., Vangay, P., Huang, H., Ward, T., Hillmann, B. M., Al-Ghalith, G. A., Travis, D. A., Long, H. T., Tuan, B. V., Minh, V. V., Cabana, F., Nadler, T., Toddes, B., Murphy, T., Glander, K. E., Johnson, T. J. & Knights, D. 2016. Captivity humanizes the primate microbiome. *Proceedings of the National Academy of Sciences of the United States of America,* 113**,** 10376-81, 10.1073/pnas.1521835113.

Clooney, A. G., Fouhy, F., Sleator, R. D., A, O. D., Stanton, C., Cotter, P. D. & Claesson, M. J. 2016. Comparing Apples and Oranges?: Next Generation Sequencing and Its Impact on Microbiome Analysis. *PLoS ONE,* 11**,** e0148028, 10.1371/journal.pone.0148028.

Colosimo, P. F., Hosemann, K. E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R. M., Schluter, D. & Kingsley, D. M. 2005. Widespread Parallel Evolution in Sticklebacks by Repeated Fixation of Ectodysplasin Alleles. *Science,* 307**,** 1928-1933.

Cordaux, R., Paces-Fessy, M., Raimond, M., Michel-Salzat, A., Zimmer, M. & Bouchon, D. 2007. Molecular characterization and evolution of arthropod-pathogenic Rickettsiella bacteria. *Applied and Environmental Microbiology,* 73**,** 5045-7, 10.1128/AEM.00378-07.

Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannan, B. J. & Relman, D. A. 2012. The application of ecological theory toward an understanding of the human microbiome. *Science,* 336**,** 1255-62, 10.1126/science.1224203.

Crawford, K., Mcdonald, R. A. & Bearhop, S. 2008. Applications of stable isotope techniques to the ecology of mammals. *Mammal Review,* 38**,** 87-107.

Crispo, E. 2007. The Baldwin effect and genetic assimilation: revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution,* 61**,** 2469-79, 10.1111/j.1558-5646.2007.00203.x.

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., Ling, A. V., Devlin, A. S., Varma, Y., Fischbach, M. A., Biddinger, S. B., Dutton, R. J. & Turnbaugh, P. J. 2014. Diet rapidly and reproducibly alters the human gut microbiome. *Nature,* 505**,** 559-63, 10.1038/nature12820.

Davidson, G. L., Raulo, A. & Knowles, S. C. L. 2020. Identifying Microbiome-Mediated Behaviour in Wild Vertebrates. *Trends in Ecology & Evolution,* 35**,** 972-980, 10.1016/j.tree.2020.06.014.

Davies, K. G. 2009. Understanding the interaction between an obligate hyperparasitic bacterium, Pasteuria penetrans and its obligate plant-parasitic nematode host, Meloidogyne spp. *Advances in Parasitology,* 6810.1016/S0065-308X(08)00609-X.

Davis, J. P., Youssef, N. H. & Elshahed, M. S. 2009. Assessment of the diversity, abundance, and ecological distribution of members of candidate division SR1 reveals a high level of phylogenetic diversity but limited morphotypic diversity. *Applied and Environmental Microbiology,* 75**,** 4139-48, 10.1128/AEM.00137-09.

De Villemereuil, P., Gaggiotti, O. E., Mouterde, M. & Till-Bottraud, I. 2016. Common garden experiments in the genomic era: new perspectives and opportunities. *Heredity (Edinb),* 116**,** 249-54, 10.1038/hdy.2015.93.

Defaveri, J. & Merila, J. 2014. Local adaptation to salinity in the three-spined stickleback? *Journal of Evolutionary Biology,* 27**,** 290-302, 10.1111/jeb.12289.

Dehler, C. E., Secombes, C. J. & Martin, S. a. M. 2017. Seawater transfer alters the intestinal microbiota profiles of Atlantic salmon (Salmo salar L.). *Sci Rep,* 7**,** 13877, 10.1038/s41598-017-13249-8.

Delsuc, F., Metcalf, J. L., Wegener Parfrey, L., Song, S. J., Gonzalez, A. & Knight, R. 2014. Convergence of gut microbiomes in myrmecophagous mammals. *Molecular Ecology,* 23**,** 1301-17, 10.1111/mec.12501.

Deniro, M. J. & Epstein, S. 1978. Influence of diet on the distribution of carbon isotopes in animals. *Geochimica et Cosmochimica Acta,* 42**,** 496 - 506.

Des Marais, D. L., Hernandez, K. M. & Juenger, T. E. 2013. Genotype-by-Environment Interaction and Plasticity: Exploring Genomic Responses of Plants to the Abiotic Environment. *Annual Review of Ecology, Evolution, and Systematics,* 44**,** 5-29, 10.1146/annurev-ecolsys-110512-135806.

Desai, A. R., Links, M. G., Collins, S. A., Mansfield, G. S., Drew, M. D., Van Kessel, A. G. & Hill, J. E. 2012. Effects of plant-based diets on the distal gut microbiome of rainbow trout (Oncorhynchus mykiss). *Aquaculture,* 350-353**,** 134-142, 10.1016/j.aquaculture.2012.04.005.

Desantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P. & Andersen, G. L. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology,* 72**,** 5069-72, 10.1128/AEM.03006-05.

Dheilly, N. M. 2014. Holobiont-Holobiont interactions: redefining host-parasite interactions. *PLoS Pathog,* 10**,** e1004093, 10.1371/journal.ppat.1004093.

Dheilly, N. M., Martinez Martinez, J., Rosario, K., Brindley, P. J., Fichorova, R. N., Kaye, J. Z., Kohl, K. D., Knoll, L. J., Lukes, J., Perkins, S. L., Poulin, R., Schriml, L. & Thompson, L. R. 2019. Parasite microbiome project: Grand challenges. *PLoS Pathog,* 15**,** e1008028, 10.1371/journal.ppat.1008028.

Dheilly, N. M., Poulin, R. & Thomas, F. 2015. Biological warfare: Microorganisms as drivers of host-parasite interactions. *Infection Genetics and Evolution,* 34**,** 251-9, 10.1016/j.meegid.2015.05.027.

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 of the United States of America,* 107**,** 11971-5, 10.1073/pnas.1002601107.

Dominianni, C., Sinha, R., Goedert, J. J., Pei, Z., Yang, L., Hayes, R. B. & Ahn, J. 2015. Sex, body mass index, and dietary fiber intake influence the human gut microbiome. *PLoS ONE,* 10**,** e0124599, 10.1371/journal.pone.0124599.

Dominianni, C., Wu, J., Hayes, R. B. & Ahn, J. 2014. Comparison of methods for fecal microbiome biospecimen collection. *BMC Microbiol,* 1410.1186/1471-2180-14-103.

Dosmann, A., Bahet, N. & Gordon, D. M. 2016. Experimental modulation of external microbiome affects nestmate recognition in harvester ants (Pogonomyrmex barbatus). *PeerJ,* 4**,** e1566, 10.7717/peerj.1566.

Douglas, A. E. 2019. Simple animal models for microbiome research. *Nat Rev Microbiol,* 17**,** 764-775, 10.1038/s41579-019-0242-1.

Ducarmon, Q. R., Zwittink, R. D., Hornung, B., Van Schaik, W., Young, V. B. & Kuijper, E. J. 2019. Gut Microbiota and Colonization Resistance against Bacterial Enteric Infection. *Microbiology and Molecular Biology Reviews,* 8310.1128/MMBR.00007-19.

Dulski, T., Kozlowski, K. & Ciesielski, S. 2020. Habitat and seasonality shape the structure of tench (Tinca tinca L.) gut microbiome. *Sci Rep,* 10**,** 4460, 10.1038/s41598-020-61351-1.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics,* 26**,** 2460-1, 10.1093/bioinformatics/btq461.

Edgar, R. C. 2017. Accuracy of microbial community diversity estimated by closed- and open-reference OTUs. *PeerJ,* 5**,** e3889, 10.7717/peerj.3889.

Eichmiller, J. J., Hamilton, M. J., Staley, C., Sadowsky, M. J. & Sorensen, P. W. 2016. Environment shapes the fecal microbiome of invasive carp species. *Microbiome,* 4**,** 44, 10.1186/s40168-016-0190-1.

Eizaguirre, C. & Lenz, T. L. 2010. Major histocompatibility complex polymorphism: dynamics and consequences of parasite-mediated local adaptation in fishes. *Journal of Fish Biology,* 77**,** 2023-47, 10.1111/j.1095-8649.2010.02819.x.

Eizaguirre, C., Lenz, T. L., Kalbe, M. & Milinski, M. 2012a. Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. *Ecology Letters,* 15**,** 723-31, 10.1111/j.1461-0248.2012.01791.x.

Eizaguirre, C., Lenz, T. L., Kalbe, M. & Milinski, M. 2012b. Rapid and adaptive evolution of MHC genes under parasite selection in experimental vertebrate populations. *Nat Commun,* 3**,** 621, 10.1038/ncomms1632.

Eizaguirre, C., Lenz, T. L., Sommerfeld, R. D., Harrod, C., Kalbe, M. & Milinski, M. 2011. Parasite diversity, patterns of MHC II variation and olfactory based mate choice in diverging three-spined stickleback ecotypes. *Evolutionary Ecology,* 25**,** 605-622, 10.1007/s10682-010-9424-z.

Eizaguirre, C., Lenz, T. L., Traulsen, A. & Milinski, M. 2009a. Speciation accelerated and stabilized by pleiotropic major histocompatibility complex immunogenes. *Ecology Letters,* 12**,** 5-12, 10.1111/j.1461-0248.2008.01247.x.

Eizaguirre, C., Yeates, S. E., Lenz, T. L., Kalbe, M. & Milinski, M. 2009b. MHC-based mate choice combines good genes and maintenance of MHC polymorphism. *Molecular Ecology,* 18**,** 3316-29, 10.1111/j.1365-294X.2009.04243.x.

Elderman, M., De Vos, P. & Faas, M. 2018a. Role of Microbiota in Sexually Dimorphic Immunity. *Front Immunol,* 9**,** 1018, 10.3389/fimmu.2018.01018.

Elderman, M., Hugenholtz, F., Belzer, C., Boekschoten, M., De Haan, B., De Vos, P. & Faas, M. 2018b. Changes in intestinal gene expression and microbiota composition during late pregnancy are mouse strain dependent. *Sci Rep,* 8**,** 10001, 10.1038/s41598-018-28292-2.

Elmer, K. R., Kusche, H., Lehtonen, T. K. & Meyer, A. 2010. Local variation and parallel evolution: morphological and genetic diversity across a species complex of neotropical crater lake cichlid fishes. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 365**,** 1763-1782, 10.1098/rstb.2009.0271.

Emde, S., Kochmann, J., Kuhn, T., Plath, M. & Klimpel, S. 2014. Getting what is served? Feeding ecology influencing parasite-host interactions in invasive round goby Neogobius melanostomus. *PLoS ONE,* 9**,** e109971, 10.1371/journal.pone.0109971.

Faith, J. J., Mcnulty, N. P., Rey, F. E. & Gordon, J. I. 2011. Predicting a Human Gut Microbiota's Response to Diet in Gnotobiotic Mice. *Science,* 333**,** 101 - 104.

Fang, B., Kemppainen, P., Momigliano, P., Feng, X. & Merilä, J. 2020. On the causes of geographically heterogeneous parallel evolution in sticklebacks. *Nature Ecology & Evolution* 4**,** 1105–15, 10.1038/s41559-020-1222-6

Ferreiro, A., Crook, N., Gasparrini, A. J. & Dantas, G. 2018. Multiscale Evolutionary Dynamics of Host-Associated Microbiomes. *Cell,* 172**,** 1216-1227, 10.1016/j.cell.2018.02.015.

Feulner, P. G., Chain, F. J., Panchal, M., Huang, Y., Eizaguirre, C., Kalbe, M., Lenz, T. L., Samonte, I. E., Stoll, M., Bornberg-Bauer, E., Reusch, T. B. & Milinski, M. 2015. Genomics of divergence along a continuum of parapatric population differentiation. *PLoS Genetics,* 11**,** e1004966, 10.1371/journal.pgen.1004966.

Fitzpatrick, B. M. 2012. Underappreciated Consequences of Phenotypic Plasticity for Ecological Speciation. *International Journal of Ecology,* 2012**,** 1-12, 10.1155/2012/256017.

Fontaine, S. S., Novarro, A. J. & Kohl, K. D. 2018. Environmental temperature alters the digestive performance and gut microbiota of a terrestrial amphibian. *Journal of Experimental Biology,* 22110.1242/jeb.187559.

Ford, S. A. & King, K. C. 2016. Harnessing the Power of Defensive Microbes: Evolutionary Implications in Nature and Disease Control. *PLoS Pathog,* 12**,** e1005465, 10.1371/journal.ppat.1005465.

Ford, S. A., Williams, D., Paterson, S. & King, K. C. 2017. Co-evolutionary dynamics between a defensive microbe and a pathogen driven by fluctuating selection. *Molecular Ecology,* 26**,** 1778-1789, 10.1111/mec.13906.

Foster, K. R. & Bell, T. 2012. Competition, not cooperation, dominates interactions among culturable microbial species. *Current Biology,* 22**,** 1845-50, 10.1016/j.cub.2012.08.005.

Foster, K. R., Schluter, J., Coyte, K. Z. & Rakoff-Nahoum, S. 2017. The evolution of the host microbiome as an ecosystem on a leash. *Nature,* 548**,** 43-51, 10.1038/nature23292.

Foti, M., Giacopello, C., Bottari, T., Fisichella, V., Rinaldo, D. & Mammina, C. 2009. Antibiotic Resistance of Gram Negatives isolates from loggerhead sea turtles (Caretta caretta) in the central Mediterranean Sea. *Marine Pollution Bulletin,* 58**,** 1363-6, 10.1016/j.marpolbul.2009.04.020.

Fraser, D. J., Weir, L. K., Bernatchez, L., Hansen, M. M. & Taylor, E. B. 2011. Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. *Heredity (Edinb),* 106**,** 404-20, 10.1038/hdy.2010.167.

Fraune, S. & Bosch, T. C. 2007. Long-term maintenance of species-specific bacterial microbiota in the basal metazoan Hydra. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 13146-51, 10.1073/pnas.0703375104.

Fredensborg, B. L., Fossdal, I. K. I., Johannesen, T. B., Stensvold, C. R., Nielsen, H. V. & Kapel, C. M. O. 2020. Parasites modulate the gut-microbiome in insects: A proof-of-concept study. *PLoS ONE,* 15**,** e0227561, 10.1371/journal.pone.0227561.

Friberg, I. M., Taylor, J. D. & Jackson, J. A. 2019. Diet in the Driving Seat: Natural Diet-Immunity-Microbiome Interactions in Wild Fish. *Front Immunol,* 10**,** 243, 10.3389/fimmu.2019.00243.

Frick, M. G., Williams, K. L., Bolten, A. B., Bjorndal, K. A. & Martins, H. R. 2009. Foraging ecology of oceanic-stage loggerhead turtles Caretta caretta. *Endangered Species Research,* 9**,** 91-97, 10.3354/esr00227.

Frischknecht, M. 1993. The breeding colouration of male sticklebacks (Gasterosteus aculeatus) energy investment in vigour three-spined as an indicator of energy investment in vigour. *Evolutionary Ecology,* 7**,** 439 - 450.

Fry, B. 2006. *Stable Isotope Ecology,* New York, Springer.

Fukuda, K., Ogawa, M., Taniguchi, H. & Saito, M. 2016. Molecular Approaches to Studying Microbial Communities: Targeting the 16S Ribosomal RNA Gene. *J UOEH,* 38**,** 223-32, 10.7888/juoeh.38.223.

Gallart, M., Adair, K. L., Love, J., Meason, D. F., Clinton, P. W., Xue, J. & Turnbull, M. H. 2018. Host Genotype and Nitrogen Form Shape the Root Microbiome of Pinus radiata. *Microbial Ecology,* 75**,** 419-433, 10.1007/s00248-017-1055-2.

Garcia-Lopez, R., Cornejo-Granados, F., Lopez-Zavala, A. A., Sanchez-Lopez, F., Cota-Huizar, A., Sotelo-Mundo, R. R., Guerrero, A., Mendoza-Vargas, A., Gomez-Gil, B. & Ochoa-Leyva, A. 2020. Doing More with Less: A Comparison of 16S Hypervariable Regions in Search of Defining the Shrimp Microbiota. *Microorganisms,* 810.3390/microorganisms8010134.

Garud, N. R. & Pollard, K. S. 2020. Population Genetics in the Human Microbiome. *Trends in Genetics,* 36**,** 53-67, 10.1016/j.tig.2019.10.010.

Gerardo, N. M. & Parker, B. J. 2014. Mechanisms of symbiont-conferred protection against natural enemies: an ecological and evolutionary framework. *Curr Opin Insect Sci,* 4**,** 8-14, 10.1016/j.cois.2014.08.002.

Gevers, D., Pop, M., Schloss, P. D. & Huttenhower, C. 2012. Bioinformatics for the Human Microbiome Project. *PLoS Comput Biol,* 8**,** e1002779, 10.1371/journal.pcbi.1002779.

Ghanbari, M., Kneifel, W. & Domig, K. J. 2015. A new view of the fish gut microbiome: Advances from next-generation sequencing. *Aquaculture,* 448**,** 464-475, 10.1016/j.aquaculture.2015.06.033.

Gibbs, P. A. 2009. Pathogenic Clostridium species. *Foodborne Pathogens.* Woodhead Publishing.

Gibson, G. 2005. The Synthesis and Evolution of a Supermodel. *Science,* 307**,** 1890-1891.

Gibson, K. M., Nguyen, B. N., Neumann, L. M., Miller, M., Buss, P., Daniels, S., Ahn, M. J., Crandall, K. A. & Pukazhenthi, B. 2019. Gut microbiome differences between wild and captive black rhinoceros - implications for rhino health. *Sci Rep,* 9**,** 7570, 10.1038/s41598-019-43875-3.

Gilbert, J. A., Jansson, J. K. & Knight, R. 2014. The Earth Microbiome project: successes and aspirations. *BMC Biology,* 1210.1186/s12915-014-0069-1.

Gilbert, J. A., Jansson, J. K. & Knight, R. 2018. Earth Microbiome Project and Global Systems Biology. *mSystems,* 310.1128/mSystems.00217-17.

Gilbert, S. F., Sapp, J. & Tauber, A. I. 2012. A symbiotic view of life: we have never been individuals. *Quarterly Review of Biology,* 87**,** 325-41, 10.1086/668166.

Giraffa, G. & Neviani, E. 2001. DNA-based, culture-independent strategies for evaluating microbial communities in food-associated ecosystems. *Internation journal of food microbiology,* 67**,** 19 - 34.

Glasl, B., Smith, C. E., Bourne, D. G. & Webster, N. S. 2019. Disentangling the effect of host-genotype and environment on the microbiome of the coral Acropora tenuis. *PeerJ,* 7**,** e6377, 10.7717/peerj.6377.

Godon, J. J., Arulazhagan, P., Steyer, J. P. & Hamelin, J. 2016. Vertebrate bacterial gut diversity: size also matters. *BMC Ecology,* 16**,** 12, 10.1186/s12898-016-0071-2.

Godon, J. J., Moriniere, J., Moletta, M., Gaillac, M., Bru, V. & Delgenes, J. P. 2005. Rarity associated with specific ecological niches in the bacterial world: the 'Synergistes' example. *Environmental Microbiology,* 7**,** 213-24, 10.1111/j.1462-2920.2004.00693.x.

Gongora, E., Elliott, K. H. & Whyte, L. 2021. Gut microbiome is affected by inter-sexual and inter-seasonal variation in diet for thick-billed murres (Uria lomvia). *Sci Rep,* 11**,** 1200, 10.1038/s41598-020-80557-x.

Gould, A. L., Zhang, V., Lamberti, L., Jones, E. W., Obadia, B., Korasidis, N., Gavryushkin, A., Carlson, J. M., Beerenwinkel, N. & Ludington, W. B. 2018. Microbiome interactions shape host fitness. *Proceedings of the National Academy of Sciences of the United States of America,* 115**,** E11951-E11960, 10.1073/pnas.1809349115.

Green, T. J., Smullen, R. & Barnes, A. C. 2013. Dietary soybean protein concentrate-induced intestinal disorder in marine farmed Atlantic salmon, Salmo salar is associated with alterations in gut microbiota. *Veterinary Microbiology,* 166**,** 286-92, 10.1016/j.vetmic.2013.05.009.

Greenblatt, R. J., Quackenbush, S. L., Casey, R. N., Rovnak, J., Balazs, G. H., Work, T. M., Casey, J. W. & Sutton, C. A. 2005. Genomic variation of the fibropapilloma-associated marine turtle herpesvirus across seven geographic areas and three host species. *J Virol,* 79**,** 1125-32, 10.1128/JVI.79.2.1125-1132.2005.

Greischar, M. A. & Koskella, B. 2007. A synthesis of experimental work on parasite local adaptation. *Ecology Letters,* 10**,** 418-34, 10.1111/j.1461-0248.2007.01028.x.

Grey, J., Thackeray, S. J., Jones, R. I. & Shine, A. 2002. Ferox Trout (Salmo trutta) as 'Russian dolls': complementary gut content and stable isotope analyses of the Loch Ness foodweb. *Freshwater Biology,* 47**,** 1235-1243.

Greyson-Gaito, C. J., Bartley, T. J., Cottenie, K., Jarvis, W. M. C., Newman, A. E. M. & Stothart, M. R. 2020. Into the wild: microbiome transplant studies need broader ecological reality. *Proc Biol Sci,* 287**,** 20192834, 10.1098/rspb.2019.2834.

Griffith, T. M. & Watson, M. A. 2005. Stress avoidance in a common annual: reproductive timing is important for local adaptation and geographic distribution. *Journal of Evolutionary Biology,* 18**,** 1601-12, 10.1111/j.1420-9101.2005.01021.x.

Grube, M., Koberl, M., Lackner, S., Berg, C. & Berg, G. 2012. Host-parasite interaction and microbiome response: effects of fungal infections on the bacterial community of the Alpine lichen Solorina crocea. *FEMS Microbiol Ecol,* 82**,** 472-81, 10.1111/j.1574-6941.2012.01425.x.

Haldane, J. B. S. 1949. Disease and Evolution. *La Ricerca Scientifica Supplemento,* 19**,** 1 - 11.

Hamilton, W. D. & Zuk, M. 1982. Heritable true fitness and bright birdsL A role for parasites? *Science,* 21810.1126/science.7123238.

Hamley, M., Franke, F., Kurtz, J. & Scharsack, J. P. 2017. An experimental approach to the immuno-modulatory basis of host-parasite local adaptation in tapeworm-infected sticklebacks. *Experimental Parasitology,* 180**,** 119-132, 10.1016/j.exppara.2017.03.004.

Hammer, T. J., Sanders, J. G. & Fierer, N. 2019. Not all animals need a microbiome. *FEMS Microbiol Lett,* 36610.1093/femsle/fnz117.

Hansen, G. H. & Olafsen, J. A. 1999. Bacterial Interactions in Early Life Stages of Marine Cold Water Fish. *Microbial Ecology,* 38**,** 1-26, 10.1007/s002489900158.

Hanson, D., Hu, J., Hendry, A. P. & Barrett, R. D. H. 2017. Heritable gene expression differences between lake and stream stickleback include both parallel and antiparallel components. *Heredity,* 119**,** 339-348, 10.1038/hdy.2017.50.

Harer, A., Torres-Dowdall, J., Rometsch, S. J., Yohannes, E., Machado-Schiaffino, G. & Meyer, A. 2020. Parallel and non-parallel changes of the gut microbiota during trophic diversification in repeated young adaptive radiations of sympatric cichlid fish. *Microbiome,* 8**,** 149, 10.1186/s40168-020-00897-8.

Hasnain, S. Z., Gallagher, A. L., Grencis, R. K. & Thornton, D. J. 2013. A new role for mucins in immunity: insights from gastrointestinal nematode infection. *International Journal of Biochemistry & Cell Biology,* 45**,** 364-74, 10.1016/j.biocel.2012.10.011.

Hawkes, L. A., Broderick, A. C., Coyne, M. S., Godfrey, M. H., Lopez-Jurado, L. F., Lopez-Suarez, P., Merino, S. E., Varo-Cruz, N. & Godley, B. J. 2006. Phenotypically linked dichotomy in sea turtle foraging requires multiple conservation approaches. *Current Biology,* 16**,** 990-5, 10.1016/j.cub.2006.03.063.

He, S., Ran, C., Qin, C., Li, S., Zhang, H., De Vos, W. M., Ringo, E. & Zhou, Z. 2017. Anti-Infective Effect of Adhesive Probiotic Lactobacillus in Fish is Correlated With Their Spatial Distribution in the Intestinal Tissue. *Sci Rep,* 7**,** 13195, 10.1038/s41598-017-13466-1.

Heinritz, S. N., Weiss, E., Eklund, M., Aumiller, T., Louis, S., Rings, A., Messner, S., Camarinha-Silva, A., Seifert, J., Bischoff, S. C. & Mosenthin, R. 2016. Intestinal Microbiota and Microbial Metabolites Are Changed in a Pig Model Fed a High-Fat/Low-Fiber or a Low-Fat/High-Fiber Diet. *PLoS ONE,* 11**,** e0154329, 10.1371/journal.pone.0154329.

Hendry, A. P., Hudson, K., Walker, J. A., Rasanen, K. & Chapman, L. J. 2011. Genetic divergence in morphology-performance mapping between Misty Lake and inlet stickleback. *Journal of Evolutionary Biology,* 24**,** 23-35, 10.1111/j.1420-9101.2010.02155.x.

Hendry, A. P., Taylor, E. B. & Mcphail, J. D. 2002. Adaptive divergence and the balance between selection and gene flow: lake and stream stickleback in the Mister system. *Evolution,* 56**,** 1199-1216.

Henry, L. P., Bruijning, M., Forsberg, S. K. G. & Ayroles, J. F. 2021. The microbiome extends host evolutionary potential. *Nat Commun,* 12**,** 5141, 10.1038/s41467-021-25315-x.

Herre, E. A., Knowlton, N., Mueller, U. G. & Rehner, S. A. 1999. The evolution of mutualisms: exploring the paths between conflict and cooperation. *Trends in Ecology & Evolution,* 14.

Hildebrandt, M. A., Hoffmann, C., Sherrill-Mix, S. A., Keilbaugh, S. A., Hamady, M., Chen, Y. Y., Knight, R., Ahima, R. S., Bushman, F. & Wu, G. D. 2009. High-fat diet determines the composition of the murine gut microbiome independently of obesity. *Gastroenterology,* 137**,** 1716-24 10.1053/j.gastro.2009.08.042.

Hird, S. M. 2017. Evolutionary Biology Needs Wild Microbiomes. *Front Microbiol,* 8**,** 1-10, 10.3389/fmicb.2017.00725.

Hird, S. M., Carstens, B. C., Cardiff, S. W., Dittmann, D. L. & Brumfield, R. T. 2014. Sampling locality is more detectable than taxonomy or ecology in the gut microbiota of the brood-parasitic Brown-headed Cowbird (Molothrus ater). *PeerJ,* 2**,** e321, 10.7717/peerj.321.

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., Poss, M. L., Reed, L. K., Storfer, A. & Whitlock, M. C. 2016. Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. *American Naturalist,* 188**,** 379-97, 10.1086/688018.

Holm, J. B., Sorobetea, D., Kiilerich, P., Ramayo-Caldas, Y., Estelle, J., Ma, T., Madsen, L., Kristiansen, K. & Svensson-Frej, M. 2015. Chronic Trichuris muris Infection Decreases Diversity of the Intestinal Microbiota and Concomitantly Increases the Abundance of Lactobacilli. *PLoS ONE,* 10**,** e0125495, 10.1371/journal.pone.0125495.

Hornef, M. 2015. Pathogens, Commensal Symbionts, and Pathobionts: Discovery and Functional Effects on the Host. *ILAR Journal,* 56**,** 159-62, 10.1093/ilar/ilv007.

Houlden, A., Hayes, K. S., Bancroft, A. J., Worthington, J. J., Wang, P., Grencis, R. K. & Roberts, I. S. 2015. Chronic Trichuris muris Infection in C57BL/6 Mice Causes Significant Changes in Host Microbiota and Metabolome: Effects Reversed by Pathogen Clearance. *PLoS ONE,* 10**,** e0125945, 10.1371/journal.pone.0125945.

Huang, Y., Chain, F. J., Panchal, M., Eizaguirre, C., Kalbe, M., Lenz, T. L., Samonte, I. E., Stoll, M., Bornberg-Bauer, E., Reusch, T. B., Milinski, M. & Feulner, P. G. 2016. Transcriptome profiling of immune tissues reveals habitat-specific gene expression between lake and river sticklebacks. *Molecular Ecology,* 25**,** 943-58, 10.1111/mec.13520.

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., Creasy, H. H., Earl, A. M., Fitzgerald, M. G., Fulton, R. S. & Giglio, M. G. 2012. Structure, function and diversity of the healthy human microbiome. *Nature,* 486**,** 207-14, 10.1038/nature11234.

Imhann, F., Bonder, M. J., Vich Vila, A., Fu, J., Mujagic, Z., Vork, L., Tigchelaar, E. F., Jankipersadsing, S. A., Cenit, M. C., Harmsen, H. J., Dijkstra, G., Franke, L., Xavier, R. J., Jonkers, D., Wijmenga, C., Weersma, R. K. & Zhernakova, A. 2016. Proton pump inhibitors affect the gut microbiome. *Gut,* 65**,** 740-8, 10.1136/gutjnl-2015-310376.

Isaacson, R. & Kim, H. B. 2012. The intestinal microbiome of the pig. *Anim Health Res Rev,* 13**,** 100-9, 10.1017/S1466252312000084.

Itoh, H., Hori, T., Sato, Y., Nagayama, A., Tago, K., Hayatsu, M. & Kikuchi, Y. 2018. Infection dynamics of insecticide-degrading symbionts from soil to insects in response to insecticide spraying. *ISME J,* 12**,** 909-920, 10.1038/s41396-017-0021-9.

Jackson, J. M., Pimsler, M. L., Oyen, K. J., Strange, J. P., Dillon, M. E. & Lozier, J. D. 2020. Local adaptation across a complex bioclimatic landscape in two montane bumble bee species. *Molecular Ecology,* 29**,** 920-939, 10.1111/mec.15376.

Jaenike, J., Unckless, R., Cockburn, S. N., Boelio, L. M. & Perlman, S. J. 2010. Adaptation via Symbiosis: Recent Spread of a Drosophila Defensive Symbiont. *Science,* 32910.1126/science.1188235.

Jervis, P., Pintanel, P., Hopkins, K., Wierzbicki, C., Shelton, J. M. G., Skelly, E., Rosa, G. M., Almeida-Reinoso, D., Eugenia-Ordonez, M., Ron, S., Harrison, X., Merino-Viteri, A. & Fisher, M. C. 2021. Post-epizootic microbiome associations across communities of neotropical amphibians. *Molecular Ecology,* 30**,** 1322-1335, 10.1111/mec.15789.

Jin, Y., Xia, J., Pan, Z., Yang, J., Wang, W. & Fu, Z. 2018. Polystyrene microplastics induce microbiota dysbiosis and inflammation in the gut of adult zebrafish. *Environmental Pollution,* 235**,** 322-329, 10.1016/j.envpol.2017.12.088.

Johnson, C. K., Tinker, M. T., Estes, J. A., Conrad, P. A., Staedler, M., Miller, M. A., Jessup, D. A. & Mazet, J. A. 2009. Prey choice and habitat use drive sea otter pathogen exposure in a resource-limited coastal system. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 2242-7, 10.1073/pnas.0806449106.

Johnson, M. W., Nelson, P. A. & Dick, T. A. 2004. Structuring mechanisms of yellow perch (Perca flavescens) parasite communities: host age, diet, and local factors. *Canadian Journal of Zoology,* 82**,** 1291-1301, 10.1139/z04-092.

Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M. C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M. C., Myers, R. M., Miller, C. T., Summers, B. R., Knecht, A. K., Brady, S. D., Zhang, H., Pollen, A. A., Howes, T., Amemiya, C., Broad Institute Genome Sequencing, P., Whole Genome Assembly, T., Baldwin, J., Bloom, T., Jaffe, D. B., Nicol, R., Wilkinson, J., Lander, E. S., Di Palma, F., Lindblad-Toh, K. & Kingsley, D. M. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature,* 484**,** 55-61, 10.1038/nature10944.

Jones, K., Ariel, E., Burgess, G. & Read, M. 2016. A review of fibropapillomatosis in Green turtles (Chelonia mydas). *Veterinary Journal,* 212**,** 48-57, 10.1016/j.tvjl.2015.10.041.

Kalbe, M., Eizaguirre, C., Dankert, I., Reusch, T. B., Sommerfeld, R. D., Wegner, K. M. & Milinski, M. 2009. Lifetime reproductive success is maximized with optimal major histocompatibility complex diversity. *Proc Biol Sci,* 276**,** 925-34, 10.1098/rspb.2008.1466.

Kalbe, M. & Kurtz, J. 2006. Local differences in immunocompetence reflect resistance of sticklebacks against the eye fluke Diplostomum pseudospathaceum. *Parasitology,* 132**,** 105-16, 10.1017/S0031182005008681.

Kalbe, M., Wegner, K. M. & Reusch, T. B. 2002. Dispersion patterns of parasites in 0+ year three-spined sticklebacks: a cross population comparison. *Journal of Fish Biology,* 60**,** 1529-1542, 10.1111/j.1095-8649.2002.tb02445.x.

Kan, H., Zhao, F., Zhang, X. X., Ren, H. & Gao, S. 2015. Correlations of Gut Microbial Community Shift with Hepatic Damage and Growth Inhibition of Carassius auratus Induced by Pentachlorophenol Exposure. *Environmental Science & Technology,* 49**,** 11894-902, 10.1021/acs.est.5b02990.

Kaufmann, J., Eizaguirre, C., Milinski, M. & Lenz, T. L. 2015. The contribution of post-copulatory mechanisms to incipient ecological speciation in sticklebacks. *Biol Lett,* 11**,** 20140933, 10.1098/rsbl.2014.0933.

Kaufmann, J., Lenz, T. L., Kalbe, M., Milinski, M. & Eizaguirre, C. 2017. A field reciprocal transplant experiment reveals asymmetric costs of migration between lake and river ecotypes of three-spined sticklebacks (Gasterosteus aculeatus). *Journal of Evolutionary Biology,* 30**,** 938 - 950.

Kawecki, T. J. & Ebert, D. 2004. Conceptual issues in local adaptation. *Ecology Letters,* 7**,** 1225-1241, 10.1111/j.1461-0248.2004.00684.x.

Keane, R. & Berleman, J. 2016. The predatory life cycle of Myxococcus xanthus. *Microbiology (Reading),* 162**,** 1-11, 10.1099/mic.0.000208.

Keenan, S. W., Engel, A. S. & Elsey, R. M. 2013. The alligator gut microbiome and implications for archosaur symbioses. *Sci Rep,* 3**,** 2877, 10.1038/srep02877.

Kelly, C. & Salinas, I. 2017. Under Pressure: Interactions between Commensal Microbiota and the Teleost Immune System. *Front Immunol,* 8**,** 559, 10.3389/fimmu.2017.00559.

Kiers, E. T., Rousseau, R. A., West, S. A. & Denison, R. F. 2003. Host sanctions and the legume–rhizobium mutualism. *Science,* 425**,** 78-81, 10.1038/nature01931.

Kim, D., Hofstaedter, C. E., Zhao, C., Mattei, L., Tanes, C., Clarke, E., Lauder, A., Sherrill-Mix, S., Chehoud, C., Kelsen, J., Conrad, M., Collman, R. G., Baldassano, R., Bushman, F. D. & Bittinger, K. 2017a. Optimizing methods and dodging pitfalls in microbiome research. *Microbiome,* 5**,** 52, 10.1186/s40168-017-0267-5.

Kim, J. J. & Khan, W. I. 2013. Goblet cells and mucins: role in innate defense in enteric infections. *Pathogens,* 2**,** 55-70, 10.3390/pathogens2010055.

Kim, M., Morrison, M. & Yu, Z. 2011. Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *Journal of Microbiological Methods,* 84**,** 81-7, 10.1016/j.mimet.2010.10.020.

Kim, S., Covington, A. & Pamer, E. G. 2017b. The intestinal microbiota: Antibiotics, colonization resistance, and enteric pathogens. *Immunological Reviews,* 279**,** 90-105, 10.1111/imr.12563.

King, K. C. & Bonsall, M. B. 2017. The evolutionary and coevolutionary consequences of defensive microbes for host-parasite interactions. *BMC Evolutionary Biology,* 17**,** 190, 10.1186/s12862-017-1030-z.

King, K. C., Brockhurst, M. A., Vasieva, O., Paterson, S., Betts, A., Ford, S. A., Frost, C. L., Horsburgh, M. J., Haldenby, S. & Hurst, G. D. 2016. Rapid evolution of microbe-mediated protection against pathogens in a worm host. *ISME J,* 10**,** 1915-24, 10.1038/ismej.2015.259.

Kleinertz, S., Klimpel, S. & Palm, H. W. 2012. Parasite communities and feeding ecology of the European sprat (Sprattus sprattus L.) over its range of distribution. *Parasitology Research,* 110**,** 1147-57, 10.1007/s00436-011-2605-z.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M. & Glockner, F. O. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research,* 41**,** e1, 10.1093/nar/gks808.

Knowles, S. C. L., Eccles, R. M. & Baltrunaite, L. 2019. Species identity dominates over environment in shaping the microbiota of small mammals. *Ecology Letters,* 10.1111/ele.13240.

Kohl, K. D., Cary, T. L., Karasov, W. H. & Dearing, M. D. 2013. Restructuring of the amphibian gut microbiota through metamorphosis. *Environ Microbiol Rep,* 5**,** 899-903, 10.1111/1758-2229.12092.

Kohl, K. D., Varner, J., Wilkening, J. L. & Dearing, M. D. 2018. Gut microbial communities of American pikas (Ochotona princeps): Evidence for phylosymbiosis and adaptations to novel diets. *Journal of Animal Ecology,* 87**,** 323-330, 10.1111/1365-2656.12692.

Kohl, K. D. & Yahn, J. 2016. Effects of environmental temperature on the gut microbial communities of tadpoles. *Environmental Microbiology,* 18**,** 1561-5, 10.1111/1462-2920.13255.

Kolodny, O. & Schulenburg, H. 2020. Microbiome-mediated plasticity directs host evolution along several distinct time scales. *Philosophical Transactions of the Royal Society of London B Biological Sciences,* 375**,** 20190589, 10.1098/rstb.2019.0589.

Koonin, E. V. & Wolf, Y. I. 2012. Evolution of microbes and viruses: a paradigm shift in evolutionary biology? *Front Cell Infect Microbiol,* 2**,** 119, 10.3389/fcimb.2012.00119.

Koren, O., Goodrich, J. K., Cullender, T. C., Spor, A., Laitinen, K., Backhed, H. K., Gonzalez, A., Werner, J. J., Angenent, L. T., Knight, R., Backhed, F., Isolauri, E., Salminen, S. & Ley, R. E. 2012. Host remodeling of the gut microbiome and metabolic changes during pregnancy. *Cell,* 150**,** 470-80, 10.1016/j.cell.2012.07.008.

Korsnes, K., Nicolaisen, O., Skår, C. K., Nerland, A. H. & Bergh, Ø. 2006. Bacteria in the gut of juvenile cod Gadus morhua fed live feed enriched with four different commercial diets. *ICES Journal of Marine Science,* 63**,** 296-301, 10.1016/j.icesjms.2005.10.012.

Koskella, B. & Bergelson, J. 2020. The study of host-microbiome (co)evolution across levels of selection. *Philosophical Transactions of the Royal Society of London B Biological Sciences,* 375**,** 20190604, 10.1098/rstb.2019.0604.

Koskella, B., Hall, L. J. & Metcalf, C. J. E. 2017. The microbiome beyond the horizon of ecological and evolutionary theory. *Nat Ecol Evol,* 1**,** 1606-1615, 10.1038/s41559-017-0340-2.

Kostic, A. D., Howitt, M. R. & Garrett, W. S. 2013. Exploring host-microbiota interactions in animal models and humans. *Genes & Development,* 27**,** 701-18, 10.1101/gad.212522.112.

Kragelund, C., Kong, Y., Van Der Waarde, J., Thelen, K., Eikelboom, D., Tandoi, V., Thomsen, T. R. & Nielsen, P. H. 2006. Ecophysiology of different filamentous Alphaproteobacteria in industrial wastewater treatment plants. *Microbiology (Reading),* 152**,** 3003-3012, 10.1099/mic.0.29249-0.

Krause, E., Wichels, A., Gimenez, L., Lunau, M., Schilhabel, M. B. & Gerdts, G. 2012. Small changes in pH have direct effects on marine bacterial community composition: a microcosm approach. *PLoS ONE,* 7**,** e47035, 10.1371/journal.pone.0047035.

Kurtz, J., Kalbe, M., Aeschlimann, P. B., Haberli, M. A., Wegner, K. M., Reusch, T. B. & Milinski, M. 2004. Major histocompatibility complex diversity influences parasite resistance and innate immunity in sticklebacks. *Proc Biol Sci,* 271**,** 197-204, 10.1098/rspb.2003.2567.

Kwiatkowski, M., Engelstadter, J. & Vorburger, C. 2012. On genetic specificity in symbiont-mediated host-parasite coevolution. *PLoS Comput Biol,* 8**,** e1002633, 10.1371/journal.pcbi.1002633.

Lafferty, K. D., Allesina, S., Arim, M., Briggs, C. J., De Leo, G., Dobson, A. P., Dunne, J. A., Johnson, P. T., Kuris, A. M., Marcogliese, D. J., Martinez, N. D., Memmott, J., Marquet, P. A., Mclaughlin, J. P., Mordecai, E. A., Pascual, M., Poulin, R. & Thieltges, D. W. 2008. Parasites in food webs: the ultimate missing links. *Ecology Letters,* 11**,** 533-46, 10.1111/j.1461-0248.2008.01174.x.

Laloë, J.-O., Cozens, J., Renom, B., Taxonera, A. & Hays, G. C. 2014. Effects of rising temperature on the viability of an important sea turtle rookery. *Nature Climate Change,* 4**,** 513-518, 10.1038/nclimate2236.

Langille, M. G., Zaneveld, J., Caporaso, J. G., Mcdonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkepile, D. E., Vega Thurber, R. L., Knight, R., Beiko, R. G. & Huttenhower, C. 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol,* 31**,** 814-21, 10.1038/nbt.2676.

Larsen, A. M., Mohammed, H. H. & Arias, C. R. 2014. Characterization of the gut microbiota of three commercially valuable warmwater fish species. *Journal of Applied Microbiology,* 116**,** 1396-404, 10.1111/jam.12475.

Laukens, D., Brinkman, B. M., Raes, J., De Vos, M. & Vandenabeele, P. 2016. Heterogeneity of the gut microbiome in mice: guidelines for optimizing experimental design. *FEMS Microbiol Rev,* 40**,** 117-32, 10.1093/femsre/fuv036.

Lawley, T. D. & Walker, A. W. 2013. Intestinal colonization resistance. *Immunology,* 138**,** 1-11, 10.1111/j.1365-2567.2012.03616.x.

Lee, S. C., Tang, M. S., Lim, Y. A., Choy, S. H., Kurtz, Z. D., Cox, L. M., Gundra, U. M., Cho, I., Bonneau, R., Blaser, M. J., Chua, K. H. & Loke, P. 2014. Helminth colonization is associated with increased diversity of the gut microbiota. *PLoS Negl Trop Dis,* 8**,** e2880, 10.1371/journal.pntd.0002880.

Lee, W., Hayakawa, T., Kurihara, Y., Hanzawa, M., Sawada, A., Kaneko, A., Morimitsu, Y., Natsume, T., Aisu, S., Ito, T., Honda, T. & Hanya, G. 2021. Stomach and colonic microbiome of wild Japanese macaques. *American Journal of Primatology,* 83**,** e23242, 10.1002/ajp.23242.

Lee, W. J. & Hase, K. 2014. Gut microbiota-generated metabolites in animal health and disease. *Nat Chem Biol,* 10**,** 416-24, 10.1038/nchembio.1535.

Lee, Y. K. & Mazmanian, S. K. 2010. Has the microbiota played a critical role in the evolution of the adaptive immune system? *Science,* 330**,** 1768-73, 10.1126/science.1195568.

Lefevre, T., Lebarbenchon, C., Gauthier-Clerc, M., Misse, D., Poulin, R. & Thomas, F. 2009. The ecological significance of manipulative parasites. *Trends in Ecology & Evolution,* 24**,** 41-8, 10.1016/j.tree.2008.08.007.

Lenz, T. L., Eizaguirre, C., Rotter, B., Kalbe, M. & Milinski, M. 2013. Exploring local immunological adaptation of two stickleback ecotypes by experimental infection and transcriptome-wide digital gene expression analysis. *Molecular Ecology,* 22**,** 774-86, 10.1111/j.1365-294X.2012.05756.x.

Leung, J. M., Graham, A. L. & Knowles, S. C. L. 2018. Parasite-Microbiota Interactions With the Vertebrate Gut: Synthesis Through an Ecological Lens. *Front Microbiol,* 9**,** 843, 10.3389/fmicb.2018.00843.

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., Schlegel, M. L., Tucker, T. A., Schrenzel, M. D. & Gordon, J. I. 2008a. Evolution of mammals and their gut microbes. *Science,* 320**,** 1647 - 1651, 10.1126/science.1155725.

Ley, R. E., Lozupone, C. A., Hamady, M., Knight, R. & Gordon, J. I. 2008b. Worlds within worlds: evolution of the vertebrate gut microbiota. *Nature Reviews Microbiology,* 610.1038/nrmicro1978.

Ley, R. E., Peterson, D. A. & Gordon, J. I. 2006. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell,* 124**,** 837-48, 10.1016/j.cell.2006.02.017.

Li, J., Ni, J., Li, J., Wang, C., Li, X., Wu, S., Zhang, T., Yu, Y. & Yan, Q. 2014. Comparative study on gastrointestinal microbiota of eight fish species with different feeding habits. *Journal of Applied Microbiology,* 117**,** 1750-60, 10.1111/jam.12663.

Li, Q., Lauber, C. L., Czarnecki-Maulden, G., Pan, Y. & Hannah, S. S. 2017. Effects of the Dietary Protein and Carbohydrate Ratio on Gut Microbiomes in Dogs of Different Body Conditions. *mBio,* 810.1128/mBio.01703-16.

Li, T., Long, M., Ji, C., Shen, Z., Gatesoupe, F. J., Zhang, X., Zhang, Q., Zhang, L., Zhao, Y., Liu, X. & Li, A. 2016. Alterations of the gut microbiome of largemouth bronze gudgeon (Coreius guichenoti) suffering from furunculosis. *Sci Rep,* 6**,** 30606, 10.1038/srep30606.

Liew, W. C., Bartfai, R., Lim, Z., Sreenivasan, R., Siegfried, K. R. & Orban, L. 2012. Polygenic sex determination system in zebrafish. *PLoS ONE,* 7**,** e34397, 10.1371/journal.pone.0034397.

Lindström, E. S. & Bergström, A. K. 2004. Influence of inlet bacteria on bacterioplankton assemblage composition in lakes of different hydraulic retention time. *Limnology and Oceanography,* 49**,** 125-136.

Ling, F., Steinel, N., Weber, J., Ma, L., Smith, C., Correa, D., Zhu, B., Bolnick, D. & Wang, G. 2020. The gut microbiota response to helminth infection depends on host sex and genotype. *ISME J,* 10.1038/s41396-020-0589-3.

Liu, H., Guo, X., Gooneratne, R., Lai, R., Zeng, C., Zhan, F. & Wang, W. 2016. The gut microbiome and degradation enzyme activity of wild freshwater fishes influenced by their trophic levels. *Sci Rep,* 6**,** 24340, 10.1038/srep24340.

Llewellyn, M. S., Boutin, S., Hoseinifar, S. H. & Derome, N. 2014. Teleost microbiomes: the state of the art in their characterization, manipulation and importance in aquaculture and fisheries. *Front Microbiol,* 5**,** 207, 10.3389/fmicb.2014.00207.

Llewellyn, M. S., Leadbeater, S., Garcia, C., Sylvain, F. E., Custodio, M., Ang, K. P., Powell, F., Carvalho, G. R., Creer, S., Elliot, J. & Derome, N. 2017. Parasitism perturbs the mucosal microbiome of Atlantic Salmon. *Sci Rep,* 7**,** 43465, 10.1038/srep43465.

Lochmiller, R. L. & Deerenberg, C. 2000. Trade-offs in evolutionary immunology: just what is the cost of immunity? *Oikos,* 8810.1034/j.1600-0706.2000.880110.x.

Locke, S. A., Marcogliese, D. J. & Valtonen, E. T. 2014. Vulnerability and diet breadth predict larval and adult parasite diversity in fish of the Bothnian Bay. *Oecologia,* 174**,** 253-62, 10.1007/s00442-013-2757-x.

Lockley, E. C., Fouda, L., Correia, S. M., Taxonera, A., Nash, L. N., Fairweather, K., Reischig, T., Durão, J., Dinis, H., Roque, S. M., Lomba, J. P., Dos Passos, L., Cameron, S. J. K., Stiebens, V. A. & Eizaguirre, C. 2020. Long-term survey of sea turtles (Caretta caretta) reveals correlations between parasite infection, feeding ecology, reproductive success and population dynamics. *Scientific Reports,* 1010.1038/s41598-020-75498-4.

Lopez-Garcia, A., Pineda-Quiroga, C., Atxaerandio, R., Perez, A., Hernandez, I., Garcia-Rodriguez, A. & Gonzalez-Recio, O. 2018. Comparison of Mothur and QIIME for the Analysis of Rumen Microbiota Composition Based on 16S rRNA Amplicon Sequences. *Front Microbiol,* 9**,** 3010, 10.3389/fmicb.2018.03010.

Lorrain, A., Paulet, Y. M., Chauvaud, L., Savoye, N., Donval, A. & Saout, C. 2002. Differential δ13C and δ15N signatures among scallop tissues: Implications for ecology and physiology. *Journal of Experimental Marine Biology and Ecology,* 275**,** 47-61.

Lozupone, C. A. & Knight, R. 2007. Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 11436-40, 10.1073/pnas.0611525104.

Lutz, P. L., Musick, J. A. & Wyneken, J. 2002. *The Biology of Sea Turtles*, CRC press.

Macpherson, A. J. & Harris, N. L. 2004. Interactions between commensal intestinal bacteria and the immune system. *Nature Reviews Immunology,* 4**,** 478 - 485, 0.1038/nri1373.

Mallott, E. K., Borries, C., Koenig, A., Amato, K. R. & Lu, A. 2020. Reproductive hormones mediate changes in the gut microbiome during pregnancy and lactation in Phayre's leaf monkeys. *Sci Rep,* 10**,** 9961, 10.1038/s41598-020-66865-2.

Mandal, S., Van Treuren, W., White, R. A., Eggesbo, M., Knight, R. & Peddada, S. D. 2015. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb Ecol Health Dis,* 26**,** 27663, 10.3402/mehd.v26.27663.

Marchinko, K. B. & Schluter, D. 2007. Parallel evolution by correlated response: lateral plate reduction in threespine stickleback. *Evolution,* 61**,** 1084-90, 10.1111/j.1558-5646.2007.00103.x.

Marco, A., Abella, E., Liria-Loza, A., Martins, S., López, O., Jiménez-Bordón, S., Medina, M., Oujo, C., Gaona, P., Godley, B. J., López-Jurado, L. F., Garner, T. & Williams, R. 2012. Abundance and exploitation of loggerhead turtles nesting in Boa Vista island, Cape Verde: the only substantial rookery in the eastern Atlantic. *Animal Conservation,* 15**,** 351-360, 10.1111/j.1469-1795.2012.00547.x.

Marcogliese, D. J. & Cone, D. K. 1997. Food webs: a plea for parasites. *Trends in Ecology & Evolution,* 12**,** 320–325.

Markle, J. G. M., Frank, D. N., Mortin-Toth, S., Robertson, C. E., Feazel, L. M., Rolle-Kampczyk, U., Von Bergen, M., Mccoy, K. D., Macpherson, A. J. & Danska, J. S. 2013. Sex Differences in the Gut Microbiome Drive Hormone-Dependent Regulation of Autoimmunity. *Science,* 339**,** 1084-1088, 10.1126/science.123352.

Marques, D. A., Lucek, K., Meier, J. I., Mwaiko, S., Wagner, C. E., Excoffier, L. & Seehausen, O. 2016. Genomics of Rapid Incipient Speciation in Sympatric Threespine Stickleback. *PLoS Genetics,* 12**,** e1005887, 10.1371/journal.pgen.1005887.

Marques, J. F., Santos, M. J., Teixeira, C. M., Batista, M. I. & Cabral, H. N. 2011. Host-parasite relationships in flatfish (Pleuronectiformes)--the relative importance of host biology, ecology and phylogeny. *Parasitology,* 138**,** 107-21, 10.1017/S0031182010001009.

Martinez Arbizu, P. 2020. pairwiseAdonis: Pairwise multilevel comparison using adonis. R package version 0.4.

Matthews, B., Marchinko, K. B., Bolnick, D. & Mazumder, A. 2010. Specialization of trophic position and habitat use by sticklebacks in an adaptive radiation. *Ecology,* 91**,** 1025-1034, 10.1890/09-0235.1.

Matthews, B. & Mazumder, A. 2005. Consequences of large temporal variability of zooplankton δ15N for modeling fish trophic position and variation. *Limnology and Oceanography, 50(5), pp.1404-1414.,* 50**,** 1404-1414, 10.4319/lo.2005.50.5.1404.

Maurice, C. F., Knowles, S. C., Ladau, J., Pollard, K. S., Fenton, A., Pedersen, A. B. & Turnbaugh, P. J. 2015. Marked seasonal variation in the wild mouse gut microbiota. *ISME J,* 9**,** 2423-34, 10.1038/ismej.2015.53.

Mcdonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., Desantis, T. Z., Probst, A., Andersen, G. L., Knight, R. & Hugenholtz, P. 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J,* 6**,** 610-8, 10.1038/ismej.2011.139.

Mcfall-Ngai, M., Hadfield, M. G., Bosch, T. C., Carey, H. V., Domazet-Loso, T., Douglas, A. E., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S. F., Hentschel, U., King, N., Kjelleberg, S., Knoll, A. H., Kremer, N., Mazmanian, S. K., Metcalf, J. L., Nealson, K., Pierce, N. E., Rawls, J. F., Reid, A., Ruby, E. G., Rumpho, M., Sanders, J. G., Tautz, D. & Wernegreen, J. J. 2013. Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** 3229-36, 10.1073/pnas.1218525110.

Mcgowin, A. E., Truong, T. M., Corbett, A. M., Bagley, D. A., Ehrhart, L. M., Bresette, M. J., Weege, S. T. & Clark, D. 2011. Genetic barcoding of marine leeches (Ozobranchus spp.) from Florida sea turtles and their divergence in host specificity. *Mol Ecol Resour,* 11**,** 271-8, 10.1111/j.1755-0998.2010.02946.x.

Mckenney, E. A., Williamson, L., Yoder, A. D., Rawls, J. F., Bilbo, S. D. & Parker, W. 2015. Alteration of the rat cecal microbiome during colonization with the helminth Hymenolepis diminuta. *Gut Microbes,* 6**,** 182-93, 10.1080/19490976.2015.1047128.

Mckenzie, V. J., Song, S. J., Delsuc, F., Prest, T. L., Oliverio, A. M., Korpita, T. M., Alexiev, A., Amato, K. R., Metcalf, J. L., Kowalewski, M., Avenant, N. L., Link, A., Di Fiore, A., Seguin-Orlando, A., Feh, C., Orlando, L., Mendelson, J. R., Sanders, J. & Knight, R. 2017. The Effects of Captivity on the Mammalian Gut Microbiome. *Integrative and Comparative Biology,* 57**,** 690-704, 10.1093/icb/icx090.

Mckinnon, J. S. & Rundle, H. D. 2002. Speciation in nature: the threespine stickleback model systems. *Trends in Ecology and Evolution,* 17**,** 480 - 488.

Mcknight, D. T., Huerlimann, R., Bower, D. S., Schwarzkopf, L., Alford, R. A., Zenger, K. R. & Jarman, S. 2019. Methods for normalizing microbiome data: An ecological perspective. *Methods in Ecology and Evolution,* 10**,** 389-400, 10.1111/2041-210x.13115.

Mcmurdie, P. J. & Holmes, S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE,* 8**,** e61217, 10.1371/journal.pone.0061217.

Mcmurdie, P. J. & Holmes, S. 2014. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Computational Biology,* 10.

Meng, X. L., Li, S., Qin, C. B., Zhu, Z. X., Hu, W. P., Yang, L. P., Lu, R. H., Li, W. J. & Nie, G. X. 2018. Intestinal microbiota and lipid metabolism responses in the common carp (Cyprinus carpio L.) following copper exposure. *Ecotoxicology and Environmental Safety,* 160**,** 257-264, 10.1016/j.ecoenv.2018.05.050.

Mideo, N. 2009. Parasite adaptations to within-host competition. *Trends in Parasitology,* 25**,** 261-8, 10.1016/j.pt.2009.03.001.

Milinski, M. 1984. Parasites determine a predator's optimal feeding strategy. *Behavioral Ecology and Sociobiology,* 15**,** 35-37.

Milinski, M. & Bakker, T. C. M. 1990. Female sticklebacks use male coloration in mate choice and hence avoid parasitized males. *Nature,* 34410.1038/344330a0.

Milligan-Myhre, K., Small, C. M., Mittge, E. K., Agarwal, M., Currey, M., Cresko, W. A. & Guillemin, K. 2016. Innate immune responses to gut microbiota differ between oceanic and freshwater threespine stickleback populations. *Dis Model Mech,* 9**,** 187-98, 10.1242/dmm.021881.

Minagawa, M. & Wada, E. 1984. Stepwise enrichment of 15N along food chains: Further evidence and the relation between δ15N and animal age. *Geochimica et Cosmochimica Acta,* 48**,** 1135-1140.

Minich, J. J., Petrus, S., Michael, J. D., Michael, T. P., Knight, R. & Allen, E. E. 2020. Temporal, Environmental, and Biological Drivers of the Mucosal Microbiome in a Wild Marine Fish, Scomber japonicus. *mSphere,* 510.1128/mSphere.00401-20.

Moeller, A. H., Caro-Quintero, A., Mjungu, D., Georgiev, A. V., Lonsdorf, E. V., Muller, M. N., Pusey, A. E., Peeters, M., Hahn, B. H. & Ochman, H. 2016a. Cospeciation of gut microbiota with hominids. *Science,* 353**,** 380 - 382.

Moeller, A. H., Foerster, S., Wilson, M. L., Pusey, A. E., Hahn, B. H. & Ochman, H. 2016b. Social behavior shapes the chimpanzee pan-microbiome. *Science advances,* 210.1126/sciadv.1500997.

Møller, A. P. 1990. Effect of haematophagous mite on the barn swallow (Hirundo rustica): A test of the Hamilton Zuk Hypothesis. *Evolution,* 44**,** 771 - 784, 10.2307/2409545.

Muegge, B. D., Kuczynski, J., Knights, D., Clemente, J. C., Gonzalez, A., Fontana, L., Henrissat, B., Knight, R. & Gordon, J. I. 2011. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science,* 332**,** 970-4, 10.1126/science.1198719.

Nayak, S. K. 2010. Role of gastrointestinal microbiota in fish. *Aquaculture Research,* 41**,** 1553-1573, 10.1111/j.1365-2109.2010.02546.x.

Nolorbe-Payahua, C. D., De Freitas, A. S., Roesch, L. F. W. & Zanette, J. 2020. Environmental contamination alters the intestinal microbial community of the livebearer killifish Phalloceros caudimaculatus. *Heliyon,* 6**,** e04190, 10.1016/j.heliyon.2020.e04190.

Nyholm, S. V. & Mcfall-Ngai, M. J. 2004. The winnowing: establishing the squid-vibrio symbiosis. *Nat Rev Microbiol,* 2**,** 632-42, 10.1038/nrmicro957.

Obeng, N., Bansept, F., Sieber, M., Traulsen, A. & Schulenburg, H. 2021. Evolution of Microbiota-Host Associations: The Microbe's Perspective. *Trends Microbiol,* 10.1016/j.tim.2021.02.005.

Ohbayashi, T., Takeshita, K., Kitagawa, W., Nikoh, N., Koga, R., Meng, X. Y., Tago, K., Hori, T., Hayatsu, M., Asano, K., Kamagata, Y., Lee, B. L., Fukatsu, T. & Kikuchi, Y. 2015. Insect's intestinal organ for symbiont sorting. *Proceedings of the National Academy of Sciences of the United States of America,* 112**,** E5179-88, 10.1073/pnas.1511454112.

Oliveira, B. C. M., Murray, M., Tseng, F. & Widmer, G. 2020. The fecal microbiota of wild and captive raptors. *Anim Microbiome,* 2**,** 15, 10.1186/s42523-020-00035-7.

Oliveira-Sequeira, T. C., David, E. B., Ribeiro, C., Guimaraes, S., Masseno, A. P., Katagiri, S. & Sequeira, J. L. 2014. Effect of Bifidobacterium animalis on mice infected with Strongyloides venezuelensis. *Revista do Instituto de Medicina Tropical de Sao Paulo,* 56**,** 105-9, 10.1590/S0036-46652014000200003.

Omeyer, L. C. M., Godley, B. J. & Broderick, A. C. 2017. Growth rates of adult sea turtles. *Endangered Species Research,* 34**,** 357-371, 10.3354/esr00862.

Pace, A., Dipineto, L., Fioretti, A. & Hochscheid, S. 2019. Loggerhead sea turtles as sentinels in the western Mediterranean: antibiotic resistance and environment-related modifications of Gram-negative bacteria. *Marine Pollution Bulletin,* 149**,** 110575, 10.1016/j.marpolbul.2019.110575.

Parada, A. E., Needham, D. M. & Fuhrman, J. A. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Microbiology,,* 18**,** 1403–1414, 10.1111/1462-2920.13023.

Parker, B. J., Barribeau, S. M., Laughton, A. M., De Roode, J. C. & Gerardo, N. M. 2011. Non-immunological defense in an evolutionary framework. *Trends in Ecology & Evolution,* 26**,** 242-8, 10.1016/j.tree.2011.02.005.

Parks, B. W., Nam, E., Org, E., Kostem, E., Norheim, F., Hui, S. T., Pan, C., Civelek, M., Rau, C. D., Bennett, B. J., Mehrabian, M., Ursell, L. K., He, A., Castellani, L. W., Zinker, B., Kirby, M., Drake, T. A., Drevon, C. A., Knight, R., Gargalovic, P., Kirchgessner, T., Eskin, E. & Lusis, A. J. 2013. Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. *Cell Metab,* 17**,** 141-52, 10.1016/j.cmet.2012.12.007.

Pascoe, E. L., Hauffe, H. C., Marchesi, J. R. & Perkins, S. E. 2017. Network analysis of gut microbiota literature: an overview of the research landscape in non-human animal studies. *ISME J,* 11**,** 2644-2651, 10.1038/ismej.2017.133.

Pegg, J., Andreou, D., Williams, C. F. & Britton, J. R. 2017. Consistent patterns of trophic niche specialization in host populations infected with a non-native copepod parasite. *Parasitology,* 144**,** 945-953, 10.1017/S0031182017000075.

Peichel, C. L., Nereng, K. S., Ohgi, K. A., Cole, B. L., Colosimo, P. F., Buerkle, C. A., Schluter, D. & Kingsley, D. M. 2001. The genetic architecture of divergence between threespine stickleback species. *Nature,* 414**,** 901.

Perlman, S. J., Hunter, M. S. & Zchori-Fein, E. 2006. The emerging diversity of Rickettsia. *Proc Biol Sci,* 273**,** 2097-106, 10.1098/rspb.2006.3541.

Perofsky, A. C., Lewis, R. J., Abondano, L. A., Di Fiore, A. & Meyers, L. A. 2017. Hierarchical social networks shape gut microbial composition in wild Verreaux's sifaka. *Proc Biol Sci,* 28410.1098/rspb.2017.2274.

Petipas, R. H., Wruck, A. C. & Geber, M. A. 2020. Microbe-mediated local adaptation to limestone barrens is context dependent. *Ecology,* 10110.1002/ecy.3092.

Phillips, C. D., Phelan, G., Dowd, S. E., Mcdonough, M. M., Ferguson, A. W., Delton Hanson, J., Siles, L., Ordonez-Garza, N., San Francisco, M. & Baker, R. J. 2012. Microbiome analysis among bats describes influences of host phylogeny, life history, physiology and geography. *Molecular Ecology,* 21**,** 2617-27, 10.1111/j.1365-294X.2012.05568.x.

Pikesley, S. K., Broderick, A. C., Cejudo, D., Coyne, M. S., Godfrey, M. H., Godley, B. J., Lopez, P., López-Jurado, L. F., Elsy Merino, S., Varo-Cruz, N., Witt, M. J. & Hawkes, L. A. 2015. Modelling the niche for a marine vertebrate: a case study incorporating behavioural plasticity, proximate threats and climate change. *Ecography,* 38**,** 803-812, 10.1111/ecog.01245.

Ponton, F., Lalubin, F., Fromont, C., Wilson, K., Behm, C. & Simpson, S. J. 2011. Hosts use altered macronutrient intake to circumvent parasite-induced reduction in fecundity. *International Journal for Parasitology,* 41**,** 43-50, 10.1016/j.ijpara.2010.06.007.

Portz, D. E., Woodley, C. M. & Cech, J. J. 2006. Stress-associated impacts of short-term holding on fishes. *Reviews in Fish Biology and Fisheries,* 16**,** 125-170, 10.1007/s11160-006-9012-z.

Post, D. M. 2002. Using stable isotopes to estimate trophic position: models, methods, and assumptions. *Ecology,* 83**,** 703 - 718.

Post, D. M., Layman, C. A., Arrington, D. A., Takimoto, G., Quattrochi, J. & Montana, C. G. 2007. Getting to the fat of the matter: models, methods and assumptions for dealing with lipids in stable isotope analyses. *Oecologia,* 152**,** 179-89, 10.1007/s00442-006-0630-x.

Poulin, R. 2011. *Evolutionary ecology of parasites*, Princeton university press.

R Core Team 2020. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rafaluk-Mohr, C., Ashby, B., Dahan, D. A. & King, K. C. 2018. Mutual fitness benefits arise during coevolution in a nematode-defensive microbe model. *Evol Lett,* 2**,** 246-256, 10.1002/evl3.58.

Ramirez, M. D., Avens, L., Seminoff, J. A., Goshe, L. R. & Heppell, S. S. 2015. Patterns of loggerhead turtle ontogenetic shifts revealed through isotopic analysis of annual skeletal growth increments. *Ecosphere,* 610.1890/es15-00255.1.

Raulo, A., Allen, B. E., Troitsky, T., Husby, A., Firth, J. A., Coulson, T. & Knowles, S. C. L. 2021. Social networks strongly predict the gut microbiota of wild mice. *ISME J,* 15**,** 2601-2613, 10.1038/s41396-021-00949-3.

Rausch, S., Held, J., Fischer, A., Heimesaat, M. M., Kuhl, A. A., Bereswill, S. & Hartmann, S. 2013. Small intestinal nematode infection of mice is associated with increased enterobacterial loads alongside the intestinal tract. *PLoS ONE,* 8**,** e74026, 10.1371/journal.pone.0074026.

Ravinet, M., Prodohl, P. A. & Harrod, C. 2013. Parallel and nonparallel ecological, morphological and genetic divergence in lake-stream stickleback from a single catchment. *Journal of Evolutionary Biology,* 26**,** 186-204, 10.1111/jeb.12049.

Reid, H. I., Treasurer, J. W., Adam, B. & Birkbeck, T. H. 2009. Analysis of bacterial populations in the gut of developing cod larvae and identification of Vibrio logei, Vibrio anguillarum and Vibrio splendidus as pathogens of cod larvae. *Aquaculture,* 288**,** 36-43, 10.1016/j.aquaculture.2008.11.022.

Reid, K., Bell, M. A. & Veeramah, K. R. 2021. Threespine Stickleback: A Model System For Evolutionary Genomics. *Annu Rev Genomics Hum Genet,* 10.1146/annurev-genom-111720-081402.

Reimchen, T. E. & Nosil, P. 2001. Ecological causes of sex-biased parasitism in threespine stickleback. *Biological Journal of the Linnean Society,* 73**,** 51-63, l0.1006/bij1.2001.0523.

Rennison, D. J., Rudman, S. M. & Schluter, D. 2019a. Parallel changes in gut microbiome composition and function during colonization, local adaptation and ecological speciation. *Proc Biol Sci,* 286**,** 20191911, 10.1098/rspb.2019.1911.

Rennison, D. J., Stuart, Y. E., Bolnick, D. I. & Peichel, C. L. 2019b. Ecological factors and morphological traits are associated with repeated genomic differentiation between lake and stream stickleback. *Philosophical Transactions of the Royal Society of London B Biological Sciences,* 374**,** 20180241, 10.1098/rstb.2018.0241.

Restivo, V. E., Kidd, K. A., Surette, M. G., Bucking, C. & Wilson, J. Y. 2021. The gut content microbiome of wild-caught rainbow darter is altered during laboratory acclimation. *Comp Biochem Physiol Part D Genomics Proteomics,* 39**,** 100835, 10.1016/j.cbd.2021.100835.

Reusch, T. B. H., Wegner, K. M. & Kalbe, M. 2001. Rapid genetic divergence in postglacial populations of threespine stickleback (Gasterosteus aculeatus): the role of habitat type, drainage and geographical proximity. *Molecular Ecology,* 10**,** 2435-2445, 10.1046/j.0962-1083.2001.01366.x.

Robinson, M. D., Mccarthy, D. J. & Smyth, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* 26**,** 139-40, 10.1093/bioinformatics/btp616.

Rodriguez, R. J., Henson, J., Van Volkenburgh, E., Hoy, M., Wright, L., Beckwith, F., Kim, Y. O. & Redman, R. S. 2008. Stress tolerance in plants via habitat-adapted symbiosis. *ISME J,* 2**,** 404-16, 10.1038/ismej.2007.106.

Rosa, B. A., Supali, T., Gankpala, L., Djuardi, Y., Sartono, E., Zhou, Y., Fischer, K., Martin, J., Tyagi, R., Bolay, F. K., Fischer, P. U., Yazdanbakhsh, M. & Mitreva, M. 2018. Differential human gut microbiome assemblages during soil-transmitted helminth infections in Indonesia and Liberia. *Microbiome,* 6**,** 33, 10.1186/s40168-018-0416-5.

Rosenberg, E. & Zilber-Rosenberg, I. 2016. Microbes Drive Evolution of Animals and Plants: the Hologenome Concept. *mBio,* 7**,** e01395, 10.1128/mBio.01395-15.

Rothschild, D., Weissbrod, O., Barkan, E., Kurilshikov, A., Korem, T., Zeevi, D., Costea, P. I., Godneva, A., Kalka, I. N., Bar, N., Shilo, S., Lador, D., Vila, A. V., Zmora, N., Pevsner-Fischer, M., Israeli, D., Kosower, N., Malka, G., Wolf, B. C., Avnit-Sagi, T., Lotan-Pompan, M., Weinberger, A., Halpern, Z., Carmi, S., Fu, J., Wijmenga, C., Zhernakova, A., Elinav, E. & Segal, E. 2018. Environment dominates over host genetics in shaping human gut microbiota. *Nature,* 555**,** 210-215, 10.1038/nature25973.

Rudman, S. M., Greenblum, S., Hughes, R. C., Rajpurohit, S., Kiratli, O., Lowder, D. B., Lemmon, S. G., Petrov, D. A., Chaston, J. M. & Schmidt, P. 2019. Microbiome composition shapes rapid genomic adaptation of Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America,* 10.1073/pnas.1907787116.

Rudnick, D. & Resh, V. 2005. Stable isotopes, mesocosms and gut content analysis demonstrate trophic differences in two invasive decapod crustacea. *Freshwater Biology,* 50**,** 1323-1336, 10.1111/j.1365-2427.2005.01398.x.

Ruiz-Rodriguez, M., Martinez-Bueno, M., Martin-Vivaldi, M., Valdivia, E. & Soler, J. J. 2013. Bacteriocins with a broader antimicrobial spectrum prevail in enterococcal symbionts isolated from the hoopoe's uropygial gland. *FEMS Microbiol Ecol,* 85**,** 495-502, 10.1111/1574-6941.12138.

Ruiz-Rodriguez, M., Valdivia, E., Soler, J. J., Martin-Vivaldi, M., Martin-Platero, A. M. & Martinez-Bueno, M. 2009. Symbiotic bacteria living in the hoopoe's uropygial gland prevent feather degradation. *Journal of Experimental Biology,* 212**,** 3621-6, 10.1242/jeb.031336.

Sachs, J. L., Skophammer, R. G. & Regus, J. U. 2011. Evolutionary transitions in bacterial symbiosis. *Proceedings of the National Academy of Sciences of the United States of America,* 108 Suppl 2**,** 10800-7, 10.1073/pnas.1100304108.

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., Turner, P., Parkhill, J., Loman, N. J. & Walker, A. W. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology,* 12**,** 87.

Sanchez-Gonzales, S., Ruiz-Campos, G. & Contreras-Balderas, S. 2001. Feeding ecology and habitat of the threespine stickleback, Gasterosteus aculeatus microcephalus, in a remnant population of northwestern Baja California, Mexico. *Ecology of Freshwater Fish,* 10**,** 191-197, 10.1034/j.1600-0633.2001.100401.x.

Sangwan, N., Xia, F. & Gilbert, J. A. 2016. Recovering complete and draft population genomes from metagenome datasets. *Microbiome,* 4**,** 8, 10.1186/s40168-016-0154-5.

Savolainen, O., Lascoux, M. & Merila, J. 2013. Ecological genomics of local adaptation. *Nature Reviews Genetics,* 14**,** 807-20, 10.1038/nrg3522.

Scheelings, T. F., Moore, R. J., Van, T. T. H., Klaassen, M. & Reina, R. D. 2020. The gut bacterial microbiota of sea turtles differs between geographically distinct populations. *Endangered Species Research,* 42**,** 95-108, 10.3354/esr01042.

Schloss, P. D. 2018. Identifying and Overcoming Threats to Reproducibility, Replicability, Robustness, and Generalizability in Microbiome Research. *mBio,* 910.1128/mBio.00525-18.

Schloss, P. D., Gevers, D. & Westcott, S. L. 2011. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS ONE,* 6**,** e27310, 10.1371/journal.pone.0027310.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J. & Weber, C. F. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology,* 75**,** 7537-41, 10.1128/AEM.01541-09.

Schluter, D. 1995. Adaptive radiation in sticklebacks: trade-offs in feeding performance and growth. *Ecology,* 76**,** 82 - 90.

Schmid-Hempel, P. 2011. *Evolutionary Parasitology: The Integrated Study of Infections, Immunology, Ecology and Genetics,* Oxford, Oxford University Press.

Schmidt, E., Mykytczuk, N. & Schulte-Hostedde, A. I. 2019. Effects of the captive and wild environment on diversity of the gut microbiome of deer mice (Peromyscus maniculatus). *ISME J,* 13**,** 1293-1305, 10.1038/s41396-019-0345-8.

Schuetzenmeister, A. & Dufey, F. 2019. VCA: Variance component analysis. R package version 1.3. 4.

Schuyler, Q. A., Wilcox, C., Townsend, K. A., Wedemeyer-Strombel, K. R., Balazs, G., Van Sebille, E. & Hardesty, B. D. 2016. Risk analysis reveals global hotspots for marine debris ingestion by sea turtles. *Glob Chang Biol,* 22**,** 567-76, 10.1111/gcb.13078.

Senko, J., Mancini, A., Seminoff, J. A. & Koch, V. 2014. Bycatch and directed harvest drive high green turtle mortality at Baja California Sur, Mexico. *Biological Conservation,* 169**,** 24-30, 10.1016/j.biocon.2013.10.017.

Sevellec, M., Pavey, S. A., Boutin, S., Filteau, M., Derome, N. & Bernatchez, L. 2014. Microbiome investigation in the ecological speciation context of lake whitefish (Coregonus clupeaformis) using next-generation sequencing. *Journal of Evolutionary Biology,* 27**,** 1029-46, 10.1111/jeb.12374.

Shamblin, B. M., Bolten, A. B., Abreu-Grobois, F. A., Bjorndal, K. A., Cardona, L., Carreras, C., Clusa, M., Monzon-Arguello, C., Nairn, C. J., Nielsen, J. T., Nel, R., Soares, L. S., Stewart, K. R., Vilaca, S. T., Turkozan, O., Yilmaz, C. & Dutton, P. H. 2014. Geographic patterns of genetic variation in a broadly distributed marine vertebrate: new insights into loggerhead turtle stock structure from expanded mitochondrial DNA sequences. *PLoS ONE,* 9**,** e85956, 10.1371/journal.pone.0085956.

Sharon, G., Segal, D., Ringo, J. M., Hefetz, A., Zilber-Rosenberg, I. & Rosenberg, E. 2010. Commensal bacteria play a role in mating preference of Drosophila melanogaster. *Proceedings of the National Academy of Sciences,* 107**,** 20051 - 20056, 10.1073/pnas.1302980110.

Shaw, A. G., Sim, K., Powell, E., Cornwell, E., Cramer, T., Mcclure, Z. E., Li, M. S. & Kroll, J. S. 2016. Latitude in sample handling and storage for infant faecal microbiota studies: the elephant in the room? *Microbiome,* 4**,** 40, 10.1186/s40168-016-0186-x.

Shi, W., Guo, Y., Xu, C., Tan, S., Miao, J., Feng, Y., Zhao, H., St Leger, R. J. & Fang, W. 2014. Unveiling the mechanism by which microsporidian parasites prevent locust swarm behavior. *Proceedings of the National Academy of Sciences of the United States of America,* 111**,** 1343-8, 10.1073/pnas.1314009111.

Shu, L., Brock, D. A., Geist, K. S., Miller, J. W., Queller, D. C., Strassmann, J. E. & Disalvo, S. 2018. Symbiont location, host fitness, and possible coadaptation in a symbiosis between social amoebae and bacteria. *Elife,* 710.7554/eLife.42660.

Singh, A., Faber-Hammond, J. J., O'rourke, C. F. & Renn, S. C. P. 2019. Gut microbial diversity increases with social rank in the African cichlid fish, Astatotilapia burtoni. *Animal Behaviour,* 152**,** 79-91, 10.1016/j.anbehav.2019.04.003.

Small, C. M., Milligan-Myhre, K., Bassham, S., Guillemin, K. & Cresko, W. A. 2017. Host Genotype and Microbiota Contribute Asymmetrically to Transcriptional Variation in the Threespine Stickleback Gut. *Genome Biol Evol,* 9**,** 504-520, 10.1093/gbe/evx014.

Smith, C. C., Snowberg, L. K., Caporaso, J. G., Knight, R. & Bolnick, D. I. 2015. Dietary input of microbes and host genetic variation shape among-population differences in stickleback gut microbiota. *ISME J,* 9**,** 2515-26, 10.1038/ismej.2015.64.

Sobel, J. M., Chen, G. F., Watt, L. R. & Schemske, D. W. 2010. The biology of speciation. *Evolution,* 64**,** 295-315, 10.1111/j.1558-5646.2009.00877.x.

Sofi, M. H., Gudi, R., Karumuthil-Melethil, S., Perez, N., Johnson, B. M. & Vasu, C. 2014. pH of drinking water influences the composition of gut microbiome and type 1 diabetes incidence. *Diabetes,* 63**,** 632-44, 10.2337/db13-0981.

Sommer, M. O. A., Dantas, G. & Church, G. M. 2009. Functional Characterization of the Antibiotic Resistance Reservoir in the Human Microflora. *Science,* 32510.1126/science.1176950.

Song, S. J., Amir, A., Metcalf, J. L., Amato, K. R., Xu, Z. Z., Humphrey, G. & Knight, R. 2016. Preservation Methods Differ in Fecal Microbiome Stability, Affecting Suitability for Field Studies. *mSystems,* 110.1128/mSystems.00021-16.

Sonnenburg, J. L., Xu, J., Leip, D. D., Chen, C. H., Westover, B. P., Weatherford, J., Buhler, J. D. & Gordon, J. I. 2005. Glycan Foraging in Vivo by an Intestine-Adapted Bacterial Symbiont. *Science,* 307**,** 1955-1959, 10.1126/science.1109051.

Spor, A., Koren, O. & Ley, R. 2011. Unravelling the effects of the environment and host genotype on the gut microbiome. *Nat Rev Microbiol,* 9**,** 279-90, 10.1038/nrmicro2540.

Star, B., Haverkamp, T. H., Jentoft, S. & Jakobsen, K. S. 2013. Next generation sequencing shows high variation of the intestinal microbial species composition in Atlantic cod caught at a single location. *BMC Microbiol,* 130.1186/1471-2180-13-248.

Steinberger, A. J. 2018. asteinberger9/seq_scripts: Release v1 (Version v1). *Zenodo,* 10.5281/zenodo.1458243.

Stephens, W. Z., Burns, A. R., Stagaman, K., Wong, S., Rawls, J. F., Guillemin, K. & Bohannan, B. J. 2016. The composition of the zebrafish intestinal microbial community varies across development. *ISME J,* 10**,** 644-54, 10.1038/ismej.2015.140.

Steury, R. A., Currey, M. C., Cresko, W. A. & Bohannan, B. J. M. 2019. Population Genetic Divergence and Environment Influence the Gut Microbiome in Oregon Threespine Stickleback. *Genes (Basel),* 1010.3390/genes10070484.

Stevens, E. J., Bates, K. A. & King, K. C. 2021. Host microbiota can facilitate pathogen infection. *PLoS Pathog,* 17**,** e1009514, 10.1371/journal.ppat.1009514.

Stewart, A., Jackson, J., Barber, I., Eizaguirre, C., Paterson, R., Van West, P., Williams, C. & Cable, J. 2017. Hook, Line and Infection: A Guide to Culturing Parasites, Establishing Infections and Assessing Immune Responses in the Three-Spined Stickleback. *Advances in Parasitology,* 98**,** 39-109, 10.1016/bs.apar.2017.07.001.

Stewart, E. J. 2012. Growing unculturable bacteria. *Journal of Bacteriology,* 194**,** 4151-60, 10.1128/JB.00345-12.

Stiebens, V. A., Merino, S. E., Chain, F. J. & Eizaguirre, C. 2013a. Evolution of MHC class I genes in the endangered loggerhead sea turtle (Caretta caretta) revealed by 454 amplicon sequencing. *BMC Biochemistry,* 1310.1186/1471-2148-13-95.

Stiebens, V. A., Merino, S. E., Roder, C., Chain, F. J., Lee, P. L. & Eizaguirre, C. 2013b. Living on the edge: how philopatry maintains adaptive potential. *Proc Biol Sci,* 280**,** 20130305, 10.1098/rspb.2013.0305.

Stutz, W. E., Lau, O. L. & Bolnick, D. I. 2014. Contrasting patterns of phenotype-dependent parasitism within and among populations of threespine stickleback. *American Naturalist,* 183**,** 810-25, 10.1086/676005.

Suchodolski, J. S. 2011. Companion animals symposium: microbes and gastrointestinal health of dogs and cats. *Journal of Animal Science,* 89**,** 1520-30, 10.2527/jas.2010-3377.

Sullam, K. E., Essinger, S. D., Lozupone, C. A., O'connor, M. P., Rosen, G. L., Knight, R., Kilham, S. S. & Russell, J. A. 2012. Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. *Molecular Ecology,* 21**,** 3363 - 3378, 10.1111/j.1365-294X.2012.05552.x.

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. *ISME J,* 9**,** 1508-22, 10.1038/ismej.2014.231.

Summers, K., Mckeon, S., Sellars, J., Keusenkothen, M., Morris, J., Gloeckner, D., Pressley, C., Price, B. & Snow, H. 2003. Parasitic exploitation as an engine of diversity. *Biol Rev Camb Philos Soc,* 78**,** 639-75, 10.1017/s146479310300616x.

Suzuki, T. A. 2017. Links between Natural Variation in the Microbiome and Host Fitness in Wild Mammals. *Integrative and Comparative Biology,* 57**,** 756-769, 10.1093/icb/icx104.

Svanback, R. & Schluter, D. 2012. Niche specialization influences adaptive phenotypic plasticity in the threespine stickleback. *American Naturalist,* 180**,** 50-9, 10.1086/666000.

Tajima, K., Nonaka, I., Higuchi, K., Takusari, N., Kurihara, M., Takenaka, A., Mitsumori, M., Kajikawa, H. & Aminov, R. I. 2007. Influence of high temperature and humidity on rumen bacterial diversity in Holstein heifers. *Anaerobe,* 13**,** 57-64, 10.1016/j.anaerobe.2006.12.001.

Taylor, E. B. & Mcphail, J. D. 2000. Historical contingency and ecological determinism interact to prime speciation in sticklebacks, Gasterosteus. *Proc Biol Sci,* 267**,** 2375-84, 10.1098/rspb.2000.1294.

Taylor, H. & Cozens, J. 2010. The effects of tourism, beachfront development and increased light pollution on nesting loggerhead turtles Caretta caretta (Linnaeus, 1758) on Sal, Cape Verde Islands. *Zoologia Caboverdiana,* 1**,** 100-111.

Theis, K. R., Dheilly, N. M., Klassen, J. L., Brucker, R. M., Baines, J. F., Bosch, T. C., Cryan, J. F., Gilbert, S. F., Goodnight, C. J., Lloyd, E. A., Sapp, J., Vandenkoornhuyse, P., Zilber-Rosenberg, I., Rosenberg, E. & Bordenstein, S. R. 2016. Getting the Hologenome Concept Right: an Eco-Evolutionary Framework for Hosts and Their Microbiomes. *mSystems,* 110.1128/mSystems.00028-16.

Tollenaere, C., Pernechele, B., Makinen, H. S., Parratt, S. R., Nemeth, M. Z., Kovacs, G. M., Kiss, L., Tack, A. J. & Laine, A. L. 2014. A hyperparasite affects the population dynamics of a wild plant pathogen. *Molecular Ecology,* 23**,** 5877-87, 10.1111/mec.12908.

Tomas, J., Aznar, F. J. & Raga, J. A. 2001. Feeding ecology of the loggerhead turtle *Caretta caretta* in the western Mediterranean. *Journal of Zoology,* 255**,** 525-532, 10.1017/S0952836901001613.

Tomillo, P. S., Saba, V. S., Piedra, R., Paladino, F. V. & Spotila, J. R. 2008. Effects of illegal harvest of eggs on the population decline of leatherback turtles in Las Baulas Marine National Park, Costa Rica. *Conservation Biology,* 22**,** 1216-24, 10.1111/j.1523-1739.2008.00987.x.

Topic Popovic, N., Strunjak-Perovic, I., Sauerborn-Klobucar, R., Barisic, J., Jadan, M., Kazazic, S., Kesner-Koren, I., Prevendar Crnic, A., Suran, J., Beer Ljubic, B., Matijatko, V. & Coz-Rakovac, R. 2017. The effects of diet supplemented with Lactobacillus rhamnosuson tissue parameters of rainbow trout, Oncorhynchus mykiss (Walbaum). *Aquaculture Research,* 48**,** 2388-2401, 10.1111/are.13074.

Traut, W. & Winking, H. 2001. Meiotic chromosomes and stages of sex chromosome evolution in fish: zebrafish, platyfish and guppy. *Chromosome Research,* 910.1023/A:1012956324417.

Trevelline, B. K., Macleod, K. J., Langkilde, T. & Kohl, K. D. 2019. Gestation alters the gut microbiota of an oviparous lizard. *FEMS Microbiol Ecol,* 9510.1093/femsec/fiz086.

Tsubokawa, D., Ishiwata, K., Goso, Y., Yokoyama, T., Kanuka, H., Ishihara, K., Nakamura, T. & Tsuji, N. 2015. Induction of Sd(a)-sialomucin and sulfated H-sulfomucin in mouse small intestinal mucosa by infection with parasitic helminth. *Experimental Parasitology,* 153**,** 165-73, 10.1016/j.exppara.2015.03.008.

Tung, J., Barreiro, L. B., Burns, M. B., Grenier, J. C., Lynch, J., Grieneisen, L. E., Altmann, J., Alberts, S. C., Blekhman, R. & Archie, E. A. 2015. Social networks predict gut microbiome composition in wild baboons. *Elife,* 410.7554/eLife.05224.

Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R. & Gordon, J. I. 2007. The human microbiome project. *Nature,* 449**,** 804-10, 10.1038/nature06244.

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**,** 1027-31, 10.1038/nature05414.

Turnbaugh, P. J., Ridaura, V. K., Faith, J. J., Rey, F. E., Knight, R. & Gordon, J. I. 2009. The Effect of Diet on the Human Gut Microbiome: A Metagenomic Analysis in Humanized Gnotobiotic Mice. *Science translational medicine,* 1**,** 1-10, 10.1126/scitranslmed.3000322.

Turnbull, P. C. B., Kramer, J. M. & Melling, J. 1990. Bacillus. *In:* Topley, W. W. C. & Wilson, G. S. (eds.) *Topley and Wilson's principles of bacteriology, virology and immunity.* 8th ed. London: Edward Arnold.

Uren Webster, T. M., Consuegra, S., Hitchings, M. & De Leaniz, C. G. 2018. Interpopulation Variation in the Atlantic Salmon Microbiome Reflects Environmental and Genetic Diversity. *Applied and Environmental Microbiology,* 84**,** 1-14, 10.1128/AEM.

Uren Webster, T. M., Rodriguez-Barreto, D., Castaldo, G., Gough, P., Consuegra, S. & Garcia De Leaniz, C. 2020. Environmental plasticity and colonisation history in the Atlantic salmon microbiome: A translocation experiment. *Molecular Ecology,* 29**,** 886-898, 10.1111/mec.15369.

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 Biol Evol,* 2**,** 53-66, 10.1093/gbe/evp057.

Vaishnava, S., Yamamoto, M., Severson, K. M., Ruhn, K. A., Yu, X., Koren, O., Ley, R., Wakeland, E. K. & Hooper, L. V. 2011. The antibacterial lectin RegIIIγ promotes the spatial segregation of microbiota and host in the intestine. *Science,* 334**,** 255-258, 10.1126/science.1209791.

Van Valen, L. 1974. Molecular evolution as predicted by natural selection. *Journal of Molecular Evolution,* 310.1007/BF01796554.

Vandeputte, D., Tito, R. Y., Vanleeuwen, R., Falony, G. & Raes, J. 2017. Practical considerations for large-scale gut microbiome studies. *FEMS Microbiol Rev,* 41**,** S154-S167, 10.1093/femsre/fux027.

Vuong, H. E., Yano, J. M., Fung, T. C. & Hsiao, E. Y. 2017. The Microbiome and Host Behavior. *Annual Review of Neuroscience,* 4010.1146/annurev-neuro-072116-031347.

Waite, D. W. & Taylor, M. W. 2014. Characterizing the avian gut microbiota: membership, driving influences, and potential function. *Front Microbiol,* 5**,** 223, 10.3389/fmicb.2014.00223.

Waite, D. W. & Taylor, M. W. 2015. Exploring the avian gut microbiota: current trends and future directions. *Front Microbiol,* 6**,** 673, 10.3389/fmicb.2015.00673.

Walter, J. & Ley, R. 2011. The human gut microbiome: ecology and recent evolutionary changes. *Annual Review of Microbiology,* 65**,** 411-29, 10.1146/annurev-micro-090110-102830.

Wang, J., Thingholm, L. B., Skieceviciene, J., Rausch, P., Kummen, M., Hov, J. R., Degenhardt, F., Heinsen, F. A., Ruhlemann, M. C., Szymczak, S., Holm, K., Esko, T., Sun, J., Pricop-Jeckstadt, M., Al-Dury, S., Bohov, P., Bethune, J., Sommer, F., Ellinghaus, D., Berge, R. K., Hubenthal, M., Koch, M., Schwarz, K., Rimbach, G., Hubbe, P., Pan, W. H., Sheibani-Tezerji, R., Hasler, R., Rosenstiel, P., D'amato, M., Cloppenborg-Schmidt, K., Kunzel, S., Laudes, M., Marschall, H. U., Lieb, W., Nothlings, U., Karlsen, T. H., Baines, J. F. & Franke, A. 2016a. Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. *Nat Genet,* 48**,** 1396-1406, 10.1038/ng.3695.

Wang, W., Zheng, S., Sharshov, K., Cao, J., Sun, H., Yang, F., Wang, X. & Li, L. 2016b. Distinctive gut microbial community structure in both the wild and farmed Swan goose (Anser cygnoides). *J Basic Microbiol,* 56**,** 1299-1307, 10.1002/jobm.201600155.

Weber, J. N., Kalbe, M., Shim, K. C., Erin, N. I., Steinel, N. C., Ma, L. & Bolnick, D. I. 2017. Resist Globally, Infect Locally: A Transcontinental Test of Adaptation by Stickleback and Their Tapeworm Parasite. *American Naturalist,* 189**,** 43-57, 10.1086/689597.

Wegner, K. M., Reusch, T. B. & Kalbe, M. 2003. Multiple parasites are driving major histocompatibility complex polymorphism in the wild. *Journal of Evolutionary Biology,* 18**,** 224-232.

Welzl, H., D'adamo, P. & Lipp, H.-P. 2001. Conditioned taste aversion as a learning and memory paradigm. *Behavioural Brain Research,* 125**,** 205-213, 10.1016/S0166-4328(01)00302-3.

Wenzel, M. A. & Piertney, S. B. 2014. Fine-scale population epigenetic structure in relation to gastrointestinal parasite load in red grouse (Lagopus lagopus scotica). *Molecular Ecology,* 23**,** 4256-73, 10.1111/mec.12833.

White, E. C., Houlden, A., Bancroft, A. J., Hayes, K. S., Goldrick, M., Grencis, R. K. & Roberts, I. S. 2018. Manipulation of host and parasite microbiotas: Survival strategies during chronic nematode infection. *Science advances,* 410.1126/sciadv.aap7399.

White, J. R., Roberts, M., Yorke, J. A. & Pop, M. 2008. Figaro: a novel statistical method for vector sequence removal. *Bioinformatics,* 24**,** 462-7, 10.1093/bioinformatics/btm632.

Willis, A. D. 2019. Rarefaction, Alpha Diversity, and Statistics. *Front Microbiol,* 10**,** 2407, 10.3389/fmicb.2019.02407.

Wrzosek, L., Miquel, S., Noordine, M. L., Bouet, S., Chevalier-Curt, M. J., Robert, V., Philippe, C., Bridonneau, C., Cherbuy, C., Robbe-Masselot, C. & Langella, P. 2013. Bacteroides thetaiotaomicron and Faecalibacterium prausnitzii influence the production of mucus glycans and the development of goblet cells in the colonic epithelium of a gnotobiotic model rodent. *BMC Biology,* 1110.1186/1741-7007-11-61.

Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N. N., Kunin, V., Goodwin, L., Wu, M., Tindall, B. J., Hooper, S. D., Pati, A., Lykidis, A., Spring, S., Anderson, I. J., D'haeseleer, P., Zemla, A., Singer, M., Lapidus, A., Nolan, M., Copeland, A., Han, C., Chen, F., Cheng, J. F., Lucas, S., Kerfeld, C., Lang, E., Gronow, S., Chain, P., Bruce, D., Rubin, E. M., Kyrpides, N. C., Klenk, H. P. & Eisen, J. A. 2009. A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature,* 462**,** 1056-60, 10.1038/nature08656.

Wu, G. D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y. Y., Keilbaugh, S. A., Bewtra, M., Knights, D., Walters, W. A., Knight, R., Sinha, R., Gilroy, E., Gupta, K., Baldassano, R., Nessel, L., Li, H., Bushman, F. D. & Lewis, J. D. 2011. Linking long-term dietary patterns with gut microbial enterotypes. *Science,* 334**,** 105-108, 10.1126/science.1208344.

Wund, M. A., Baker, J. A., Clancy, B., Golub, J. L. & Foster, S. A. 2008. A test of the "flexible stem" model of evolution: ancestral plasticity, genetic accommodation, and morphological divergence in the threespine stickleback radiation. *American Naturalist,* 172**,** 449-62, 10.1086/590966.

Wund, M. A., Valena, S., Wood, S. & Baker, J. A. 2012. Ancestral plasticity and allometry in threespine stickleback reveal phenotypes associated with derived, freshwater ecotypes. *Biological Journal of the Linnean Society,* 105**,** 573-583, 10.1111/j.1095-8312.2011.01815.x.

Xu, L., Xiang, M., Zhu, W., Zhang, M., Chen, H., Huang, J., Chen, Y., Chang, Q., Jiang, J. & Zhu, L. 2020. The Behavior of Amphibians Shapes Their Symbiotic Microbiomes. *mSystems,* 510.1128/mSystems.00626-20.

Yan, Q., Li, J., Yu, Y., Wang, J., He, Z., Van Nostrand, J. D., Kempher, M. L., Wu, L., Wang, Y., Liao, L., Li, X., Wu, S., Ni, J., Wang, C. & Zhou, J. 2016. Environmental filtering decreases with fish development for the assembly of gut microbiota. *Environmental Microbiology,* 18**,** 4739-4754, 10.1111/1462-2920.13365.

Youngblut, N. D., Reischer, G. H., Walters, W., Schuster, N., Walzer, C., Stalder, G., Ley, R. E. & Farnleitner, A. H. 2019. Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. *Nat Commun,* 10**,** 2200, 10.1038/s41467-019-10191-3.

Zaneveld, J., Turnbaugh, P. J., Lozupone, C., Ley, R. E., Hamady, M., Gordon, J. I. & Knight, R. 2008. Host-bacterial coevolution and the search for new drug targets. *Current opinion in chemical biology,* 12**,** 109-114, 10.1016/j.cbpa.2008.01.015.

Zhang, Z., Hyun, J. E., Thiesen, A., Park, H., Hotte, N., Watanabe, H., Higashiyama, T. & Madsen, K. L. 2020. Sex-Specific Differences in the Gut Microbiome in Response to Dietary Fiber Supplementation in IL-10-Deficient Mice. *Nutrients,* 1210.3390/nu12072088.

Zhou, L., Limbu, S. M., Qiao, F., Du, Z. Y. & Zhang, M. 2018. Influence of Long-Term Feeding Antibiotics on the Gut Health of Zebrafish. *Zebrafish,* 15**,** 340-348, 10.1089/zeb.2017.1526.

Zhou, Z., Yao, B., Romero, J., Waines, P., Ringø, E., Emery, M., Liles, M. R. & L., M. D. 2014. Methodological approaches used to assess fish gastrointestinal communities." Aquaculture nutrition: Gut health, probiotics and prebiotics (2014): 101-127. *In:* Merrifield, D. L. & Ringø, E. (eds.) *Aquaculture Nutrition: Gut Health, Probiotics and Prebiotics,*. Hoboken, NJ: JohnWiley & Sons Ltd.

Zhu, L., Liao, R., Wu, N., Zhu, G. & Yang, C. 2019. Heat stress mediates changes in fecal microbiome and functional pathways of laying hens. *Appl Microbiol Biotechnol,* 103**,** 461-472, 10.1007/s00253-018-9465-8.

Zoetendal, E. G., Akkermans, A. D. L., Akkermans-Van Vliet, W. M., De Visser, J. a. G. M. & De Vos, W. M. 2001. The Host Genotype Affects the Bacterial Community in the Human Gastrointestinal Tract. *Microbial Ecology in Health and Disease,* 13**,** 129-134, 10.1080/089106001750462669.

Supplementary Materials – Chapter 2

Supplementary Information 2.1

Data pre-processing details:

Samples in run: 384 – not all related to this study

Number of lanes: 1

Sequence type: 300 bp paired-end read, Illumina MiSeq V3

Amplicon type: Bacteria 16S (341F-785R)

Sequencing company: LGC Genomics

Samples were demultiplexed using Illumina's CASAVA software. Illumina TruSeq™ adapters were clipped and all reads >100 bases were retained. Reads were then sorted by amplicon inline barcodes; no mismatched barcodes were kept. Barcodes were clipped post sorting. A total of 2 mismatches were allowed per primer, with primer-dimers resulting in outer primer copies being clipped from sequence. Pairs of primers were required in the sequence fragments and put into the forward-reverse orientation. Forward and reverse reads were combined using FLASh 1.2.4, with a minimum overlap of 10 bases and a maximum mismatch of 25%.

Supplementary Table 2.1 Table of good's coverage for each individual fish sample

Sample ID	Good's Coverage (%)
341F.785R.P5.A01	0.98
341F.785R.P5.A03	0.98
341F.785R.P5.A04	0.98
341F.785R.P5.A07	0.99
341F.785R.P5.A09	0.97
341F.785R.P5.A10	0.98
341F.785R.P5.A12	0.99
341F.785R.P5.B01	0.98
341F.785R.P5.B02	0.98
341F.785R.P5.B03	0.97
341F.785R.P5.B05	0.99
341F.785R.P5.B06	0.99
341F.785R.P5.B07	0.98
341F.785R.P5.B08	0.98
341F.785R.P5.B09	0.97
341F.785R.P5.B10	0.98
341F.785R.P5.B11	0.99
341F.785R.P5.B12	0.98
341F.785R.P5.C01	0.97
341F.785R.P5.C02	0.97
341F.785R.P5.C03	0.99
341F.785R.P5.C04	0.99

341F.785R.P5.C05	0.98
341F.785R.P5.C06	0.98
341F.785R.P5.C07	0.99
341F.785R.P5.C08	0.99
341F.785R.P5.C09	0.98
341F.785R.P5.C10	0.98
341F.785R.P5.C11	0.98
341F.785R.P5.C12	0.99
341F.785R.P5.D01	0.99
341F.785R.P5.D02	0.97
341F.785R.P5.D03	0.99
341F.785R.P5.D04	0.98
341F.785R.P5.D05	0.98
341F.785R.P5.D06	0.98
341F.785R.P5.D07	0.99
341F.785R.P5.D08	0.98
341F.785R.P5.D09	0.98
341F.785R.P5.D10	0.98
341F.785R.P5.D11	0.97
341F.785R.P5.D12	0.97
341F.785R.P5.E04	0.99
341F.785R.P5.E06	0.97
341F.785R.P5.E07	0.99
341F.785R.P5.E09	0.99
341F.785R.P5.E10	0.99
341F.785R.P5.E12	0.99
341F.785R.P5.F01	0.98
341F.785R.P5.F02	0.98
341F.785R.P5.F03	0.99
341F.785R.P5.F04	0.98
341F.785R.P5.F05	0.99
341F.785R.P5.F06	0.99
341F.785R.P5.F07	0.97
341F.785R.P5.F08	0.97
341F.785R.P5.F09	0.99
341F.785R.P5.F10	0.98
341F.785R.P5.F11	0.99
341F.785R.P5.F12	0.97
341F.785R.P5.G01	0.99
341F.785R.P5.G02	0.98
341F.785R.P5.G03	0.98
341F.785R.P5.G04	0.98
341F.785R.P5.G05	0.99
341F.785R.P5.G06	0.98
341F.785R.P5.G07	0.97
341F.785R.P5.G08	0.98
341F.785R.P5.G09	0.97
341F.785R.P5.G10	0.98
341F.785R.P5.G11	0.99
341F.785R.P5.G12	0.99
341F.785R.P5.H01	0.97
341F.785R.P5.H02	0.99
341F.785R.P5.H03	0.99
341F.785R.P5.H04	0.98
341F.785R.P5.H05	0.99

341F.785R.P5.H06	0.99
341F.785R.P5.H07	0.98
341F.785R.P5.H08	0.98
341F.785R.P5.H09	0.98
341F.785R.P5.H10	0.99
341F.785R.P5.H11	0.99
341F.785R.P5.H12	0.98
341F.785R.P6.A01	0.98
341F.785R.P6.A03	0.99
341F.785R.P6.A04	0.98
341F.785R.P6.A05	1.00
341F.785R.P6.A06	0.98
341F.785R.P6.A09	0.99
341F.785R.P6.A10	0.98
341F.785R.P6.B01	0.98
341F.785R.P6.B02	0.99
341F.785R.P6.B03	0.99
341F.785R.P6.B04	0.99
341F.785R.P6.B05	0.98
341F.785R.P6.B06	0.98
341F.785R.P6.B07	0.97
341F.785R.P6.B09	0.99
341F.785R.P6.B10	0.98
341F.785R.P6.B11	0.98
341F.785R.P6.C01	0.99
341F.785R.P6.C02	0.98
341F.785R.P6.C03	0.99
341F.785R.P6.C04	0.98
341F.785R.P6.C05	0.97
341F.785R.P6.C06	0.98
341F.785R.P6.C08	0.98
341F.785R.P6.C09	0.99
341F.785R.P6.C10	0.97
341F.785R.P6.D01	0.99
341F.785R.P6.D02	0.99
341F.785R.P6.D03	0.97
341F.785R.P6.D04	0.97
341F.785R.P6.D05	0.98
341F.785R.P6.D06	0.99
341F.785R.P6.D07	0.99
341F.785R.P6.D08	0.98
341F.785R.P6.D09	0.99
341F.785R.P6.D10	0.98
341F.785R.P6.E01	0.98
341F.785R.P6.E02	0.99
341F.785R.P6.E03	0.98
341F.785R.P6.E05	0.97
341F.785R.P6.E07	0.99
341F.785R.P6.E09	0.99
341F.785R.P6.E10	0.99
341F.785R.P6.F01	0.98
341F.785R.P6.F02	0.98
341F.785R.P6.F03	0.98
341F.785R.P6.F04	0.97
341F.785R.P6.F05	0.98

341F.785R.P6.F06	0.98
341F.785R.P6.F07	0.98
341F.785R.P6.F08	0.98
341F.785R.P6.F09	0.96
341F.785R.P6.F10	0.99
341F.785R.P6.G01	0.97
341F.785R.P6.G02	0.98
341F.785R.P6.G03	0.98
341F.785R.P6.G04	0.97
341F.785R.P6.G05	0.99
341F.785R.P6.G06	0.97
341F.785R.P6.G07	0.99
341F.785R.P6.G08	0.98
341F.785R.P6.G09	0.99
341F.785R.P6.G10	0.97
341F.785R.P6.H01	0.98
341F.785R.P6.H02	0.98
341F.785R.P6.H03	0.97
341F.785R.P6.H04	0.97
341F.785R.P6.H05	0.98
341F.785R.P6.H06	0.98
341F.785R.P6.H07	0.97
341F.785R.P6.H08	0.98
341F.785R.P6.H09	0.98
341F.785R.P6.H10	0.98

Supplementary Table 2.2: Mean relative abundance (%) and standard deviation of the five most dominant phyla observed across fish samples.

Phylum	Mean Relative Abundance (%)	Minimum Relative Abundance (%)	Maximum Relative Abundance (%)
Actinobacteria	48.63 ± 6.20	0.03	74.08
Proteobacteria	19.36 ± 3.14	0.03	57.55
Firmicutes	19.06 ± 4.05	0.03	78.08
Spirochaetes	6.47 ± 11.32	0.03	57.26
Tenericutes	1.52 ± 11.17	0.03	64.36
Cyanobacteria	1.51 ± 2.76	0.03	24.25
Planctomycetes	1.15 ± 1.48	0.03	9.16
Armatimonadetes	0.69 ± 10.75	0.03	46.83
Bacteroidetes	0.62 ± 1.88	0.03	11.38
Verrucomicrobia	0.54 ± 2.11	0.03	13.08
[Thermi]	0.28 ± 4.39	0.03	18.39
Chloroflexi	0.10 ± 1.07	0.04	3.86
Synergistetes	0.04 ± 3.17	0.05	5.62
Fusobacteria	0.02 ± 0.79	0.04	2.20
Acidobacteria	0.01 ± 0.24	0.04	0.55
Gemmatimonadetes	0.00 ± 0.01	0.03	0.05

Supplementary Table 2.3: Results of the linear models linking Shannon diversity index with OTU level variation for 99% similarity clustering

	d.f.	Sum Sq	Mean Sq	F value	P-value
Continent	1	0.67	0.67	3.12	0.079
Sex	1	0.08	0.08	0.37	0.542
Continent:Sex	1	0.43	0.43	2.03	0.157
Residuals	139	29.67	0.21		

Supplementary Table 2.4: Nested PERMANOVA of β-diversity metrics showing the effect of continent, standard length classes, ecotype and population nested within continent at OTU level for 99% similarity clustering. Permutations: 1000.

		Bray-Curtis			Weighted UniFrac		
	d.f.	F	R2	P-value	F	R2	P-value
Continent	(1,142)	1.94	0.01	0.080	1.01	0.01	0.089
Standard Length (Grouped)	(2,142)	1.61	0.02	0.104	1.33	0.02	0.203
Ecotype	(2,142)	1.19	0.02	0.267	1.38	0.02	0.201
Continent: Population	(7,142)	1.30	0.06	0.140	1.33	0.06	0.135
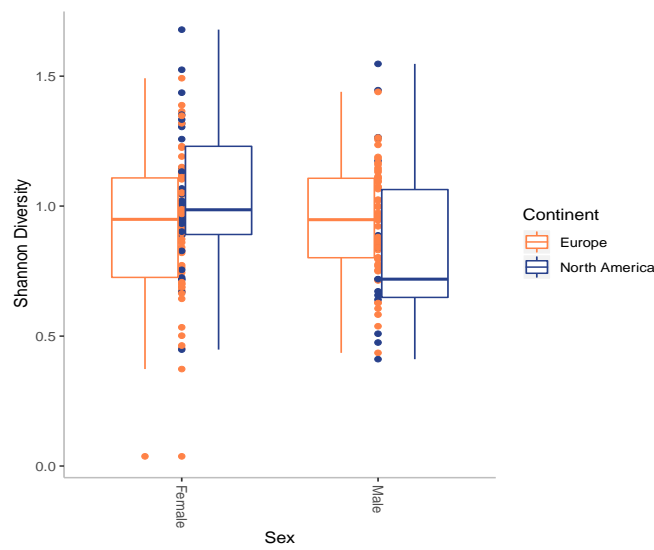
*Supplementary Figure 2.1: Shannon diversity index of phylum split by fish sex and continent.*

Supplementary Table 2.5: Results from pairwise PERMANOVA of three β-diversity metrics explained by ecotype. Permutations:1000, Bonferroni corrected. Significant results are shown in bold.

	Bray-Curtis		Weighted UniFrac		Unweighted UniFrac	
	Lake	Marine	Lake	Marine	Lake	Marine
OTU level						
Marine	**0.012**	-	**0.006**	-	**0.003**	-
River	0.614	0.072	1.000	**0.030**	**0.021**	**0.003**
Phylum level						
Marine	**0.048**	-	**0.030**	-	**0.012**	-
River	1.000	0.126	1.000	0.140	0.156	**0.003**

Supplementary Table 2.6: PERMANOVA of three β-diversity metrics for 4 lake-river ecotype pairs showing the effect of ecotype, at OTU and phylum level. Permutations:1000. Significant results are shown in bold.

	Df	*F*	R^2	*P-value*	
BRA - MIL R					
OTU level					
Bray-Curtis	(1,18)	2.11	0.11	0.062	.
Weighted UniFrac	**(1,18)**	**3.49**	**0.16**	**0.020**	*
Unweighted UniFrac	**(1,18)**	**2.58**	**0.13**	**0.002**	**
Phylum level					
Bray-Curtis	**(1,18)**	**4.48**	**0.21**	**0.024**	*
Weighted UniFrac	**(1,18)**	**4.13**	**0.20**	**0.034**	*
Unweighted UniFrac	**(1,18)**	**4.47**	**0.21**	**0.003**	**
SKO L - SKO R					
OTU level					
Bray-Curtis	**(1,39)**	**2.04**	**0.05**	**0.040**	*
Weighted UniFrac	**(1,39)**	**3.45**	**0.08**	**0.005**	**
Unweighted UniFrac	(1,39)	1.52	0.04	0.072	.
Phylum level					
Bray-Curtis	**(1,39)**	**4.62**	**0.11**	**0.007**	**
Weighted UniFrac	**(1,39)**	**4.30**	**0.10**	**0.010**	**
Unweighted UniFrac	(1,39)	0.55	0.01	0.662	
GPS - MAL					
OTU level					
Bray-Curtis	(1,27)	1.33	0.05	0.207	
Weighted UniFrac	(1,27)	1.19	0.04	0.275	
Unweighted UniFrac	(1,27)	1.19	0.04	0.471	
Phylum level					
Bray-Curtis	(1,27)	0.78	0.03	0.477	
Weighted UniFrac	(1,27)	0.80	0.03	0.520	
Unweighted UniFrac	(1,27)	0.81	0.03	0.502	
WES - EID					
OTU level					
Bray-Curtis	**(1,28)**	**2.40**	**0.08**	**0.020**	*

Weighted UniFrac	**(1,28)**	**3.18**	**0.11**	**0.014**	*****
Unweighted UniFrac	**(1,28)**	**2.60**	**0.09**	**<0.001**	*******
Phylum level					
Bray-Curtis	(1,28)	2.60	0.09	0.070	.
Weighted UniFrac	**(1,28)**	**3.26**	**0.11**	**0.037**	*****
Unweighted UniFrac	(1,28)	0.63	0.02	0.704	
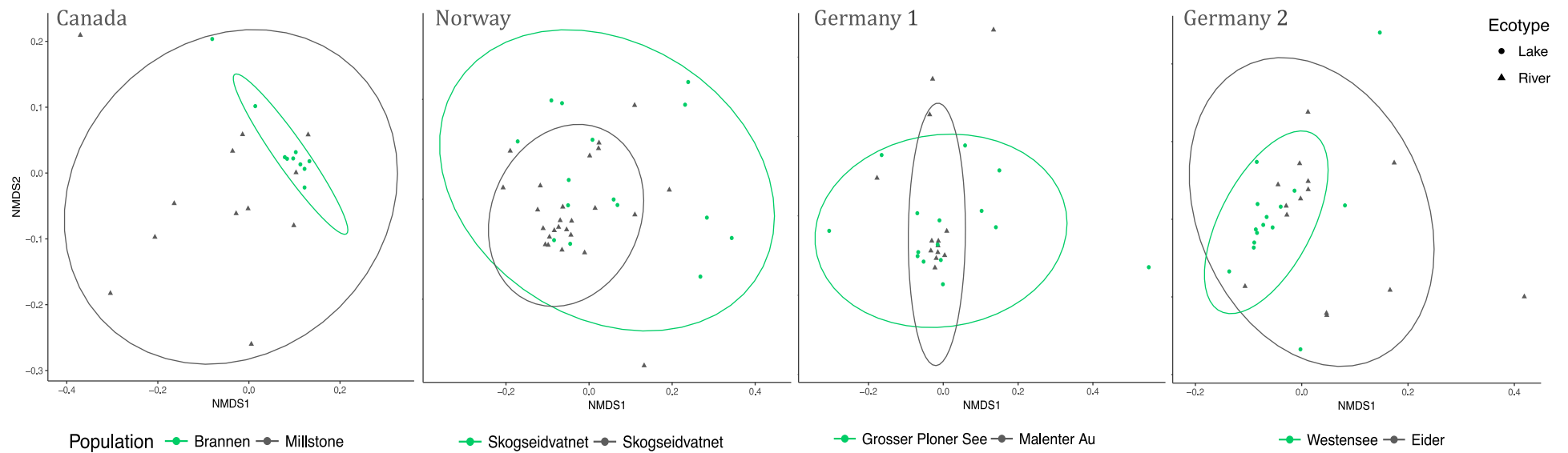
Supplementary Figure 2.2: OTU composition of microbiomes of stickleback's intestines calculated using weighted UniFrac distance. NMDS plots were subsetted into four different lake-river pairs, Brannen – Millstone River, Skogseidvatnet Lake – Skogseidvatnet River, Großer Plöner See - Malenter Au and Westensee – Eider, with 95% confidence ellipses to show each population within that ecotype. Plots coloured by population and the shape specifies ecotype.

Supplementary Materials – Chapter 3

Supplementary Table 3.1: Pairwise survival differential results for Country of origin * Habitat of exposure using TukeyHSD post hoc tests. GPS: Großer Plöner See exposure habitat, MAU: Malenter Au exposure habitat, LAB: Laboratory control habitat. Significant comparisons in bold. Df denotes degrees of freedom

Pairwise comparison	estimate	SE	Df	z.ratio	p.value
Canadian MAU - German MAU	-0.64	0.64	Inf	-1.01	0.915
Canadian MAU - Canadian GPS	**3.09**	**0.66**	**Inf**	**4.67**	**<0.001**
Canadian MAU - German GPS	-0.79	0.65	Inf	-1.22	0.827
Canadian MAU - Canadian LAB	-2.23	0.89	Inf	-2.52	0.119
Canadian MAU - German LAB	-0.74	0.78	Inf	-0.95	0.935
German MAU - Canadian GPS	**3.74**	**0.67**	**Inf**	**5.56**	**<0.001**
German MAU - German GPS	-0.15	0.65	Inf	-0.22	1.000
German MAU - Canadian LAB	-1.59	0.89	Inf	-1.78	0.478
German MAU - German LAB	-0.09	0.79	Inf	-0.12	1.000
Canadian GPS - German GPS	**-3.88**	**0.68**	**Inf**	**-5.68**	**<0.001**
Canadian GPS - Canadian LAB	**-5.32**	**0.91**	**Inf**	**-5.84**	**<0.001**
Canadian GPS - German LAB	**-3.83**	**0.81**	**Inf**	**-4.75**	**<0.001**
German GPS - Canadian LAB	-1.44	0.90	Inf	-1.61	0.594
German GPS - German LAB	0.05	0.79	Inf	0.06	1.000
Canadian LAB - German LAB	1.49	1.00	Inf	1.49	0.668

Supplementary Table 3.2: Table summarising parasite means, SE, SD, max and prevalence (% of infected hosts) across all fish sampled split by habitat of exposure

Parasite Species	Lake Exposure					River Exposure				
	Mean	*SE*	*SD*	*Max*	*Prevalence*	*Mean*	*SE*	*SD*	*Max*	*Prevalence*
Diplostomum sp.	22.18	2.32	26.46	151	99.23	0.21	0.03	0.44	2	19.23
Cyathocotyle prussica	5.23	0.35	4.04	20	93.08	0.05	0.02	0.24	2	4.81
Glochidia	5.49	0.57	6.45	42	90.00	0.06	0.02	0.29	3	4.81
Echinochasmus sp.	1.45	0.14	1.60	9	60.77	0.00	0.00	0.00	0	0.00
Gyrodactylus sp.	6.06	1.21	13.81	116	55.38	13.78	2.49	35.97	397	79.81
Camallanus lacustris	1.38	0.18	2.01	10	49.23	0.00	0.00	0.00	0	0.00
Argulus foliaceus	0.92	0.13	1.47	11	45.38	0.00	0.00	0.00	0	0.00
Tylodelphis calvata	1.15	0.23	2.62	15	31.54	0.00	0.00	0.00	0	0.00
Contracaecum sp.	0.44	0.07	0.76	4	31.54	0.00	0.00	0.00	0	0.00
Apatemon cobitis	0.25	0.05	0.61	4	18.46	0.03	0.01	0.17	1	2.88
Raphidascaris acus	0.08	0.03	0.30	2	6.92	0.03	0.02	0.23	2	2.40
Phyllodistomum folium	0.05	0.02	0.21	1	4.62	0.00	0.00	0.00	0	0.00
Proteocephalus filicollis	0.08	0.03	0.37	2	4.62	0.00	0.00	0.00	0	0.00
Nematode - eustrongo	0.03	0.02	0.17	1	3.08	0.00	0.00	0.00	0	0.00
Anguillicoloides crassus	0.02	0.02	0.20	2	1.54	0.00	0.00	0.00	0	0.00
Trematode - petasiger	0.02	0.02	0.20	2	1.54	0.00	0.00	0.00	0	0.00
Cestode - *triaenophorus*	0.02	0.02	0.26	3	0.77	0.00	0.00	0.00	0	0.00
Apiosoma sp.	0.00	0.00	0.00	0	0.00	0.96	0.68	9.78	100	0.96
Acanthocephalus lucii	0.00	0.00	0.00	0	0.00	2.90	0.23	3.27	17	78.85

Supplementary Table 3.3: ANOVA results from model parasite load explained by the interaction between fish habitat of exposure (habitat), country of origin (country), ecotype and sex. Mesocosm ID and family group were set as random factors. Significant variables in bold.

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
Habitat	**3266.2**	**3266.2**	**1**	**24.88**	**191.43**	**0.00**
Country	2.5	2.5	1	29.03	0.15	0.70
Ecotype	**118.2**	**118.2**	**1**	**48.3**	**6.93**	**0.01**
Sex	**89.1**	**89.1**	**1**	**316.47**	**5.22**	**0.02**
Habitat:Country	6.3	6.3	1	24.9	0.37	0.55
Habitat:Ecotype	**80.7**	**80.7**	**1**	**322.7**	**4.73**	**0.03**
Country:Ecotype	0.1	0.1	1	48.31	0.00	0.95
Habitat:Country:Ecotype	**89.2**	**89.2**	**1**	**322.68**	**5.23**	**0.02**

Supplementary Table 3.4: Pairwise comparisons of Parasite load (I_{PI}) for the interaction between habitat of exposure, country of origin and fish ecotype using TukeyHSD post hoc tests. Significant comparisons in bold.

Pairwise comparisons	Estimate	Std.Error	Df	t value	p. value
GPS canada Lake - GPS canada River	-4.74	2.12	249.6	-2.23	0.3359
GPS canada Lake - GPS germany Lake	-1.31	1.75	87.2	-0.75	0.9951
GPS canada Lake - GPS germany River	-3.14	1.74	85.4	-1.80	0.6228
GPS canada River - GPS germany River	1.60	1.89	92.8	0.85	0.9897
GPS germany Lake - GPS canada River	-3.42	1.90	94.4	-1.81	0.6184
GPS germany Lake - GPS germany River	-1.82	0.97	31.7	-1.89	0.5677
GPS canada Lake - MAU canada Lake	**9.00**	**1.66**	**96.1**	**5.43**	**< 0.001**
GPS canada Lake - MAU canada River	**9.78**	**1.78**	**88.6**	**5.50**	**< 0.001**
GPS canada Lake - MAU germany Lake	**9.53**	**1.74**	**86.2**	**5.46**	**< 0.001**
GPS canada Lake - MAU germany River	**7.57**	**1.75**	**86.4**	**4.34**	**0.001**
GPS canada River - MAU canada River	**14.51**	**1.84**	**98.8**	**7.89**	**< 0.001**
GPS canada River - MAU germany River	**12.30**	**1.89**	**93.7**	**6.51**	**< 0.001**
GPS germany Lake - MAU canada River	**11.09**	**1.26**	**31.6**	**8.78**	**< 0.001**
GPS germany Lake - MAU germany Lake	**10.84**	**1.08**	**27.5**	**10.07**	**< 0.001**
GPS germany Lake - MAU germany River	**8.88**	**1.22**	**29**	**7.30**	**< 0.001**
GPS germany River - MAU germany River	**10.70**	**1.06**	**26.1**	**10.07**	**< 0.001**
MAU canada Lake - GPS canada River	**-13.74**	**1.90**	**95.3**	**-7.22**	**< 0.001**
MAU canada Lake - GPS germany Lake	**-10.31**	**1.24**	**30.5**	**-8.34**	**< 0.001**
MAU canada Lake - GPS germany River	**-12.14**	**1.22**	**29.1**	**-9.92**	**< 0.001**
MAU canada River - GPS germany River	**-12.91**	**1.25**	**30.4**	**-10.33**	**< 0.001**
MAU germany Lake - GPS canada River	**-14.26**	**1.89**	**93.5**	**-7.55**	**< 0.001**
MAU germany Lake - GPS germany River	**-12.66**	**1.20**	**27.5**	**-10.55**	**< 0.001**
MAU canada Lake - MAU canada River	0.78	1.04	39.6	0.75	0.9948
MAU canada Lake - MAU germany Lake	0.53	1.22	29.3	0.43	0.9998
MAU canada Lake - MAU germany River	-1.44	1.22	29.6	-1.17	0.9334
MAU canada River - MAU germany River	-2.21	1.26	31	-1.76	0.6492
MAU germany Lake - MAU canada River	0.25	1.25	30.8	0.20	1
MAU germany Lake - MAU germany River	-1.96	0.96	31.4	-2.04	0.4741

Supplementary Table 3.5: Pairwise comparisons of I_{PI} results for country of origin and fish ecotype. Significant comparisons in bold.

Pairwise comparisons	Estimate	Std.Error	Df	t value	p. value
GPS Lake - MAU Lake	**9.92**	**0.99**	**62**	**10.03**	**< 0.001**
GPS Lake - GPS River	**-3.28**	**1.17**	**166**	**-2.82**	**0.028**
GPS Lake - MAU River	**9.33**	**1.08**	**58.5**	**8.67**	**< 0.001**
MAU Lake - GPS River	**-13.20**	**1.13**	**62.8**	**-11.74**	**< 0.001**
MAU Lake - MAU River	-0.59	0.71	35.5	-0.84	0.837
GPS River - MAU River	**12.61**	**1.06**	**65.8**	**11.88**	**< 0.001**

Supplementary Table 3.6: Pairwise PERMANOVA results comparing parasite communities between habitat of exposure, country of origin and ecotype pairs. MAU – Malenter Au, GPS – Großer Plöner See. Significant comparisons in bold.

| Pairwise comparisons	Df	Sums Of Sqs	F Model	R2	p.value	adjusted p.value
MAU_Canadian_Lake vs MAU_Canadian_River	1	0.20	4.02	0.04	0.025	0.7
MAU_Canadian_Lake vs MAU_German_Lake	1	0.08	1.39	0.01	0.229	1
MAU_Canadian_Lake vs MAU_German_River	1	0.16	2.52	0.02	0.072	1
MAU_Canadian_Lake vs GPS_Canadian_Lake	**1**	**1.70**	**26.65**	**0.31**	**0.001**	**0.028**
MAU_Canadian_Lake vs GPS_Canadian_River	**1**	**1.87**	**28.59**	**0.34**	**0.001**	**0.028**
MAU_Canadian_Lake vs GPS_German_River	**1**	**6.31**	**86.42**	**0.45**	**0.001**	**0.028**
MAU_Canadian_Lake vs GPS_German_Lake	**1**	**2.99**	**51.53**	**0.34**	**0.001**	**0.028**
MAU_Canadian_River vs MAU_German_Lake	1	0.22	6.64	0.06	0.005	0.14
MAU_Canadian_River vs MAU_German_River	**1**	**0.53**	**12.20**	**0.11**	**0.001**	**0.028**
MAU_Canadian_River vs GPS_Canadian_Lake	**1**	**1.61**	**68.02**	**0.56**	**0.001**	**0.028**
MAU_Canadian_River vs GPS_Canadian_River	**1**	**1.87**	**77.40**	**0.60**	**0.001**	**0.028**
MAU_Canadian_River vs GPS_German_River	**1**	**6.46**	**122.84**	**0.55**	**0.001**	**0.028**
MAU_Canadian_River vs GPS_German_Lake	**1**	**2.74**	**76.74**	**0.44**	**0.001**	**0.028**
MAU_German_Lake vs MAU_German_River	1	0.15	2.98	0.03	0.044	1
MAU_German_Lake vs GPS_Canadian_Lake	**1**	**1.90**	**50.13**	**0.44**	**0.001**	**0.028**
MAU_German_Lake vs GPS_Canadian_River	**1**	**2.04**	**52.74**	**0.46**	**0.001**	**0.028**
MAU_German_Lake vs GPS_German_River	**1**	**7.08**	**122.69**	**0.52**	**0.001**	**0.028**
MAU_German_Lake vs GPS_German_Lake	**1**	**3.45**	**80.34**	**0.43**	**0.001**	**0.028**
MAU_German_River vs GPS_Canadian_Lake	**1**	**2.01**	**37.39**	**0.37**	**0.001**	**0.028**
MAU_German_River vs GPS_Canadian_River	**1**	**2.08**	**37.71**	**0.38**	**0.001**	**0.028**
MAU_German_River vs GPS_German_River	**1**	**6.87**	**102.73**	**0.48**	**0.001**	**0.028**
MAU_German_River vs GPS_German_Lake	**1**	**3.72**	**70.76**	**0.40**	**0.001**	**0.028**
GPS_Canadian_Lake vs GPS_Canadian_River	1	0.09	4.18	0.21	0.009	0.252
GPS_Canadian_Lake vs GPS_German_River	1	0.39	5.78	0.08	0.004	0.112
GPS_Canadian_Lake vs GPS_German_Lake	**1**	**0.79**	**18.85**	**0.24**	**0.001**	**0.028**
GPS_Canadian_River vs GPS_German_River	1	0.24	3.54	0.05	0.012	0.336
GPS_Canadian_River vs GPS_German_Lake	**1**	**1.00**	**23.23**	**0.28**	**0.001**	**0.028**
GPS_German_River vs GPS_German_Lake	**1**	**2.30**	**37.99**	**0.26**	**0.001**	**0.028**

Supplementary Table 3.7: Pairwise PERMANOVA results comparing parasite communities between country of origin and ecotype pairs within in separated Großer Plöner See and Malenter Au datasets. Significant comparisons in bold.

Pairwise comparisons	Df	Sums Of Sqs	F Model	R2	p.value	adjusted p.value
Großer Plöner See						
Canadian Lake vs Canadian River	1	0.01102942	2.216679	0.12168406	0.043	0.258
Canadian Lake vs German River	**1**	**0.07262983**	**7.075313**	**0.09551513**	**0.001**	**0.006**
Canadian Lake vs German Lake	**1**	**0.13343131**	**19.761247**	**0.24468724**	**0.001**	**0.006**
Canadian River vs German River	**1**	**0.05693184**	**5.343589**	**0.07596412**	**0.001**	**0.006**
Canadian River vs German Lake	**1**	**0.15556984**	**22.03031**	**0.2718774**	**0.001**	**0.006**
German River vs German Lake	**1**	**0.28599157**	**30.697853**	**0.21818281**	**0.001**	**0.006**
Malenter Au						
Canadian Lake vs Canadian River	1	0.016813708	3.6628212	0.037892761	0.039	0.234
Canadian Lake vs German Lake	1	0.003054497	0.6577803	0.006225574	0.498	1
Canadian Lake vs German River	1	0.010859109	1.8821817	0.01777619	0.137	0.822
Canadian River vs German Lake	1	0.01990919	7.3184677	0.068193926	0.002	0.012
Canadian River vs German River	**1**	**0.039199719**	**10.0936076**	**0.092522447**	**0.001**	**0.006**
German Lake vs German River	1	0.011759184	2.9321601	0.025736018	0.038	0.228

Supplementary Table 3.8: Anova results of parasite load and fish fitness proxies

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
Length Growth Rate						
I$_{PI}$ Residuals	**64.32**	**64.32**	**1**	**309.21**	**6.79**	**0.010**
Location	**85.59**	**85.59**	**1**	**21.46**	**9.04**	**0.007**
Country	21.52	21.52	1	26.92	2.27	0.143
Ecotype	**203.19**	**203.19**	**1**	**25.46**	**21.46**	**< 0.001**
I$_{PI}$ Residuals:Country	**51.52**	**51.52**	**1**	**309.72**	**5.44**	**0.020**
Location:Country	47.67	47.67	1	21.49	5.03	0.036
Location:Ecotype	**125.14**	**125.14**	**1**	**300.65**	**13.21**	**< 0.001**
Country:Ecotype	**53.90**	**53.90**	**1**	**25.39**	**5.69**	**0.025**
SSI						
I$_{PI}$ Residuals	0.35	0.35	1	320.15	0.88	0.350
Location	0.11	0.11	1	314.44	0.27	0.605
Country	13.36	13.36	1	27.96	33.86	< 0.001
Ecotype	0.18	0.18	1	27.84	0.45	0.508
I$_{PI}$ Residuals:Location	0.93	0.93	1	321.79	2.35	0.126
Location:Country	0.00	0.00	1	314.81	0.01	0.924
I$_{PI}$ Residuals:Ecotype	**3.17**	**3.17**	**1**	**320.15**	**8.03**	**0.005**
Location:Ecotype	0.25	0.25	1	314.44	0.65	0.422
Country:Ecotype	0.14	0.14	1	27.96	0.36	0.551
I$_{PI}$ Residuals:Location:Ecotype	**4.12**	**4.12**	**1**	**321.79**	**10.45**	**0.001**
Location:Country:Ecotype	**2.70**	**2.70**	**1**	**314.81**	**6.85**	**0.009**

Supplementary Materials – Chapter 4
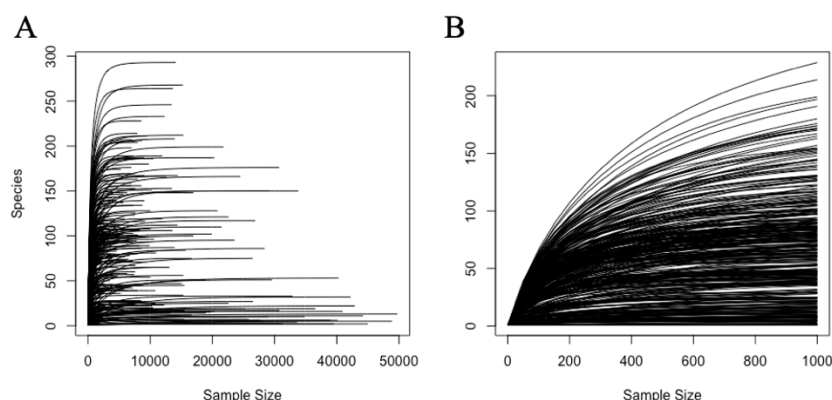
Supplementary Figure 4.1: A) Rarefaction curve for all microbial samples. B) Rarefaction curve for microbial samples rarefied to 1000 reads.

Rarefied Results

Samples were rarefied to 1000 reads and analysed using the same methods as non-rarefied analysis

Supplementary Table 4.1: Variance componence analysis on rarefied data explaining variability in fish microbial diversity due to month of collection, habitat of exposure, fish country of origin, ecotype, family group and mesocosm ID.

	DF	SS	MS	VC	%Total	SD	CV[%]
Shannon							
total	16			2.95	100.00	1.72	63.11
Month of Collection	2	130.73	65.37	0.83	28.26	0.91	33.55
Habitat of Exposure	1	22.54	22.54	0.14	4.82	0.38	13.85
Country	1	25.32	25.32	0.15	5.19	0.39	14.37
Ecotype	1	0.83	0.83	0*	0*	0*	0*
Mesocosm ID	35	154.86	4.42	0.55	18.50	0.74	27.14
error	207	263.75	1.27	1.27	43.24	1.13	41.50
Phylogenetic Diversity							
total	19			28.69	100.00	5.36	63.63
Month of Collection	2	1166.89	583.45	7.41	25.82	2.72	32.33
Habitat of Exposure	1	138.44	138.44	0.67	2.33	0.82	9.71
Country	1	276.94	276.94	1.74	6.05	1.32	15.66
Ecotype	1	22.86	22.86	0.03	0.12	0.18	2.19
Mesocosm ID	35	1487.57	42.50	4.97	17.31	2.23	26.47
Family	21	327.28	15.58	0.21	0.73	0.46	5.44
error	186	2542.84	13.67	13.67	47.65	3.70	43.92
Gini-Simpson							
total	33			0.13	100.00	0.36	50.81
Month of Collection	2	3.43	1.72	0.02	15.29	0.14	19.86
Habitat of Exposure	1	1.01	1.01	0.01	4.79	0.08	11.12
Country	1	1.38	1.38	0.01	6.68	0.09	13.13
Ecotype	1	0.02	0.02	0*	0*	0*	0*
Mesocosm ID	35	7.15	0.20	0.02	16.87	0.15	20.87
Family	21	1.55	0.07	0*	0*	0*	0*
error	186	13.92	0.07	0.07	56.37	0.27	38.14

Core microbiome results for each habitat of exposure

ASVs were classified as core microbiome if they were present in at least 65% of fish within a habitat of exposure.

Supplementary Table 4.2: ASVs identified as core within each habitat of exposure. Core ASVs were found in >65% of individuals within a habitat.

ASV	Phylum	Class	Order	Family	Genus	Species
Malenter Au Core						
81d84d2d88d5ff44cd74f85f9293cc11	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	*NA*	*NA*
707362560253e9f11f35fcd8156efbfe	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	*Ralstonia*	*NA*
cab81200dc8a1b7011cb421b1df82262	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	*Ralstonia*	*NA*
Großer Plöner See Core						
3f5225a8b2e87a448aab5f44a4412c2a	Planctomycetes	Planctomycetia	Gemmatales	Isosphaeraceae	*NA*	*NA*
0938f5e4e046f002176f9baabcbe2491	Cyanobacteria	Synechococcophycideae	Synechococcales	Synechococcaceae	*Synechococcus*	*NA*
81d84d2d88d5ff44cd74f85f9293cc11	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	*NA*	*NA*
707362560253e9f11f35fcd8156efbfe	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	*Ralstonia*	*NA*
47a55f7ba97286602de71a9eab0f5c1b	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
Laboratory Core						
4676d3a4334894acb39ce26df717597f	Planctomycetes	Planctomycetia	Pirellulales	Pirellulaceae	*NA*	*NA*
df141ac2043c0c246968f324f4116c04	Firmicutes	Bacilli	Bacillales	Bacillaceae	*NA*	*NA*
3d8a9447929371aa614dc6431bb869d3	Firmicutes	Bacilli	Bacillales	Bacillaceae	*NA*	*NA*
81d84d2d88d5ff44cd74f85f9293cc11	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	*NA*	*NA*
707362560253e9f11f35fcd8156efbfe	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	*Ralstonia*	*NA*
090f66af5d0d20a9ba480d6d24121ccb	Proteobacteria	Alphaproteobacteria	Rhizobiales	*NA*	*NA*	*NA*
7cbb74423aba3a902a330264bc88e902	Proteobacteria	Alphaproteobacteria	Rhizobiales	*NA*	*NA*	*NA*

Supplementary Table 4.3: Summary of linear mixed effect models carried out on the rarefied dataset split by the four levels of analysis – All sample types, wild fish only, Malenter Au habitat of exposure (MAU) and Großer Plöner See habitat of exposure (GPS). Shannon Diversity, Phylogenetic Diversity and Gini-Simpson all used as indices of microbial diversity as the response variable. – All variables were dropped for Großer Plöner See model so not displayed here.

| Dataset	Factor	Df	Shannon Diversity Index				Phylogenetic Diversity				Gini-Simpson			
			Sum Sq	Mean Sq	F-value	Pr(>F)	Sum Sq	Mean Sq	F-value	Pr(>F)	Sum Sq	Mean Sq	F-value	Pr(>F)
All sample types	Sample Type	2,555	20.29	10.14	6.99	0.001	259.11	129.56	8.71	<0.001	0.86	0.43	5.66	<0.001
	Habitat	2,354	10.01	5.00	3.45	0.033	1.64	0.82	0.06	0.946	0.49	0.24	3.21	0.041
	Sample Type:Habitat	4,354	46.22	11.55	7.97	<0.001	206.97	51.74	3.48	<0.001	3.05	0.76	10.01	<0.001
Wild fish	Habitat	1,25	9.02	9.02	7.19	0.013	45.98	45.98	3.41	0.076	0.26	0.26	3.66	0.065
	Country	1,27	3.67	3.67	2.93	0.098	31.55	31.55	2.34	0.138	0.29	0.29	4.04	0.053
	Ecotype	1,215.00	1.37	1.37	1.09	0.297	2.06	2.06	0.15	0.696	0.02	0.02	0.33	0.569
	I_{PI}	1,226	1.27	1.27	1.02	0.315	52.20	52.20	3.87	0.051	0.03	0.03	0.39	0.532
	Sex	1,218	1.53	1.53	1.22	0.271	7.74	7.74	0.57	0.449	0.16	0.16	2.26	0.134
	Habitat:Country	1,28	0.27	0.27	0.22	0.645	22.98	22.98	1.71	0.202	0.00	0.00	0.01	0.907
	Habitat:Ecotype	1,215	0.87	0.87	0.69	0.405	-	-	-	-	0.03	0.03	0.45	0.504
	Country:Ecotype	1,214	3.41	3.41	2.72	0.101	-	-	-	-	0.23	0.23	3.17	0.076
	Habitat:I_{PI}	1,229	0.23	0.23	0.19	0.666	26.65	26.65	1.98	0.161	0.05	0.05	0.67	0.414
	Country:I_{PI}	1,225	0.02	0.02	0.01	0.911	8.97	8.97	0.67	0.416	0.02	0.02	0.35	0.557
	Ecotype:I_{PI}	1,228	0.45	0.45	0.36	0.549	80.13	80.13	5.95	0.016	0.02	0.02	0.24	0.625
	I_{PI}:Sex	1,179	9.80	9.80	7.82	0.006	67.96	67.96	5.04	0.026	0.30	0.30	4.16	0.043
	Habitat:Country:Ecotype	1,215	5.54	5.54	4.42	0.037	-	-	-	-	0.71	0.71	9.80	0.002
	Habitat:Country:I_{PI}	1,223	1.30	1.30	1.04	0.309	73.68	73.68	5.47	0.021	0.29	0.29	3.98	0.047
	Habitat:Ecotype:I_{PI}	1,228	0.27	0.27	0.22	0.642	-	-	-	-	0.09	0.09	1.32	0.252
	Country:Ecotype:I_{PI}	1,228	0.01	0.01	0.01	0.926	-	-	-	-	0.00	0.00	0.01	0.936

Dataset	Factor	Df												
	Habitat:Country:Ecotype:I_{PI}	1,227	5.31	5.31	4.24	0.041	-	-	-	-	0.02	0.02	0.34	0.560
MAU	Country	1,23	1.92	1.92	3.34	0.081	53.10	53.10	4.82	0.037	0.13	0.13	4.14	0.053
	Ecotype	1,130	0.13	0.13	0.23	0.634	1.36	1.36	0.12	0.726	0.26	0.26	8.41	0.004
	I_{PI} MAU	1,146	9.12	9.12	15.86	0.000	173.99	173.99	15.80	0.000	0.02	0.02	0.54	0.462
	Sex	1,130	0.09	0.09	0.16	0.691	3.50	3.50	0.32	0.574	0.24	0.24	7.79	0.006
	Country:I_{PI} MAU	1,146	8.49	8.49	14.77	0.000	140.72	140.72	12.78	0.000	0.00	0.00	0.16	0.690
	Ecotype:I_{PI} MAU	1,142	3.69	3.69	6.42	0.012	74.77	74.77	6.79	0.010	0.42	0.42	13.31	0.000
	I_{PI} MAU:Sex	1,139	6.45	6.45	11.23	0.001	89.10	89.10	8.09	0.005	0.19	0.19	6.08	0.015

Supplementary Table 4.4: Summary of PERMANOVA on the rarefied dataset split by the four levels of analysis – All sample types, wild fish only, Malenter Au habitat of exposure (MAU) and Großer Plöner See habitat of exposure (GPS).

Dataset	Factor	Df	Unweighted UniFrac Distance			Weighted UniFrac Distance			Bray-Curtis Dissimilarity		
			F-value	R2	P-value	F-value	R2	P-value	F-value	R2	P-value
All sample types	Sample_Type	2	19.02	0.09	0.001	27.83	0.12	0.001	12.98	0.06	0.001
	Habitat	2	15.90	0.07	0.001	13.86	0.06	0.001	19.04	0.09	0.001
	Sample_Type:Habitat	4	4.32	0.04	0.001	5.93	0.05	0.001	5.51	0.05	0.001
	Residuals	355	-	0.80	-	-	0.77	-	-	0.80	-
	Total	363	-	1.00	-	-	1.00	-	-	1.00	-
Wild fish	Habitat	1	14.97	0.05	0.001	14.17	0.05	0.001	23.01	0.08	0.001
	Country	1	5.85	0.02	0.001	10.20	0.04	0.001	5.76	0.02	0.001
	Ecotype	1	0.84	0.00	0.685	0.80	0.00	0.640	0.81	0.00	0.742

		df	F	R²	p	F	R²	p	F	R²	p
	I_{PI} group	2	3.88	0.03	0.047	5.08	0.04	0.129	2.29	0.02	0.406
	Habitat:Country	1	3.87	0.01	0.002	5.10	0.02	0.003	4.95	0.02	0.001
	Habitat:Ecotype	1	1.11	0.00	0.242	0.53	0.00	0.824	0.81	0.00	0.703
	Country:Ecotype	1	1.72	0.01	0.046	1.47	0.01	0.150	1.38	0.00	0.126
	Habitat:I_{PI} group	2	1.40	0.01	0.088	1.36	0.01	0.228	1.22	0.01	0.156
	Country:I_{PI} group	2	1.40	0.01	0.076	1.61	0.01	0.086	1.24	0.01	0.203
	Ecotype:I_{PI} group	2	1.49	0.01	0.106	1.56	0.01	0.188	1.29	0.01	0.163
	Habitat:Country:Ecotype	1	1.02	0.00	0.404	1.72	0.01	0.105	1.52	0.01	0.064
	Habitat:Country:I_{PI} group	1	0.98	0.00	0.508	2.16	0.01	0.066	1.45	0.01	0.096
	Habitat:Ecotype:I_{PI} group	1	1.24	0.00	0.199	1.71	0.01	0.107	1.48	0.01	0.066
	Country:Ecotype:I_{PI} group	2	1.26	0.01	0.148	1.94	0.01	0.028	1.42	0.01	0.032
	Residuals	228	-	0.82	-	-	0.79	-	-	0.80	-
	Total	247	-	1.00	-	-	1.00	-	-		-
MAU	Country	1	8.19	0.05	0.001	11.90	0.07	0.001	7.45	0.05	0.001
	Ecotype	1	0.99	0.01	0.332	0.71	0.00	0.653	0.94	0.01	0.441
	I_{PI} group MAU	2	4.17	0.05	0.375	5.58	0.06	0.624	2.65	0.03	0.794
	Country:Ecotype	1	1.41	0.01	0.064	1.50	0.01	0.102	1.13	0.01	0.190
	Country:I_{PI} group MAU	2	1.04	0.01	0.263	1.80	0.02	0.029	1.19	0.01	0.078
	Ecotype:I_{PI} group MAU	2	0.81	0.01	0.787	0.77	0.01	0.678	0.98	0.01	0.489
	Country:Ecotype:I_{PI} group MAU	2	1.40	0.02	0.118	1.90	0.02	0.065	1.17	0.01	0.215
	Residuals	143	-	0.85	-	-	0.81	-	-	0.87	-
	Total	154	-	1.00	-	-	1.00	-	-	1.00	-
GPS	Country	1	1.18	0.01	0.371	3.13	0.03	0.037	3.07	0.03	0.016
	Ecotype	1	0.90	0.01	0.527	0.52	0.01	0.788	0.61	0.01	0.915
	I_{PI} group GPS	2	1.16	0.03	0.703	1.12	0.02	0.723	1.33	0.03	0.885

	df	F	R^2	p	F	R^2	p	F	R^2	p
Country:Ecotype	1	1.70	0.02	0.101	2.34	0.03	0.085	2.21	0.02	0.043
Country:I_{PI} group GPS	2	0.62	0.01	0.942	0.48	0.01	0.884	0.55	0.01	0.961
Ecotype:I_{PI} group GPS	2	0.66	0.01	0.950	0.91	0.02	0.520	1.06	0.02	0.474
Country:Ecotype:I_{PI} group GPS	2	0.78	0.02	0.855	0.63	0.01	0.888	1.05	0.02	0.662
Residuals	81	-	0.89	-	-	0.87	-	-	0.85	-
Total	92	-	1.00	-	-	1.00	-	-	1.00	-
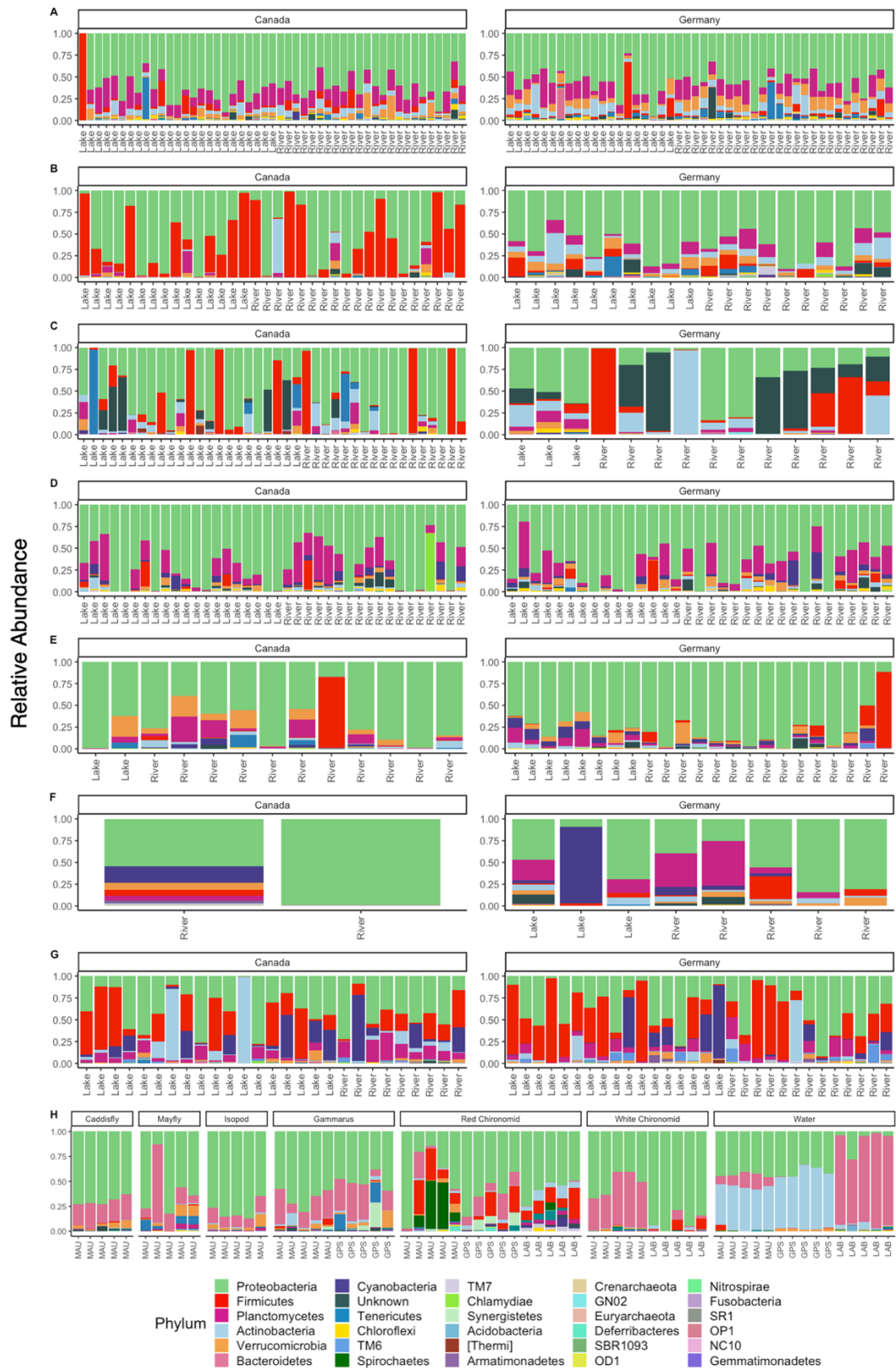
Supplementary Figure 4.2: Composition of phyla per A) Malenter Au exposure fish collected in December, B) Malenter Au exposure fish collected in May, C) Malenter Au exposure fish collected in July, D) Großer Plöner See exposure fish collected in December, E) Großer Plöner See exposure fish collected in May, F) Großer Plöner See exposure fish collected in July, G) Control exposure fish from laboratory collected in July, H) Invertebrate and water samples across all habitats of exposure collected in July.

Supplementary Table 4.5: Survival rate of fish across different month of collection, habitats of exposure, country and ecotype. Großer Plöner See – lake treatments, Malenter Au – river treatment.

Month of Collection	Habitat of Exposure	Country of Origin	Ecotype	Female Survival (%)	Male Survival (%)	Total Survival (%)
December	Großer Plöner See	Canada	Lake	88.89	88.89	88.89
			River	88.89	94.44	91.67
		Germany	Lake	100	100	100
			River	89.47	94.44	91.89
	Malenter Au	Canada	Lake	88.89	94.74	91.89
			River	88.89	100	94.59
		Germany	Lake	88.89	83.33	86.11
			River	100	95.24	97.44
May	Großer Plöner See	Canada	Lake	72.22	83.33	77.78
			River	94.74	72.22	83.78
		Germany	Lake	77.78	66.67	72.22
			River	94.44	100	97.3
	Malenter Au	Canada	Lake	100	94.44	97.22
			River	88.89	94.44	91.67
		Germany	Lake	88.89	94.44	91.67
			River	88.89	94.44	91.67
July	Großer Plöner See	Canada	Lake	13.89	13.89	13.89
			River	11.11	11.11	11.11
		Germany	Lake	91.67	59.46	75.34
			River	94.44	67.57	80.82
	Malenter Au	Canada	Lake	61.11	77.78	69.44
			River	72.97	55.56	64.38
		Germany	Lake	77.78	78.38	78.08
			River	75.68	80.56	78.08
	Lab	Canada	Lake	95	88.89	92.11
			River	100	94.44	97.22
		Germany	Lake	77.78	72.22	75
			River	100	77.78	88.89

Supplementary Table 4.6: Tukey multiple comparisons of likelihood of survival among fish country of origin and habitat of exposure. Significant results are highlighted in bold, Df denotes degrees of freedom.

Country-Habitat pair	estimate	SE	Df	z.ratio	p.value
Canada MAU - Germany MAU	-0.25	0.52	Inf	-0.48	0.965
Canada MAU - Canada GPS	**2.03**	**0.51**	**Inf**	**3.99**	**< 0.001**
Canada MAU - Germany GPS	-0.28	0.52	Inf	-0.53	0.952
Germany MAU - Canada GPS	**2.28**	**0.51**	**Inf**	**4.46**	**< 0.001**
Germany MAU - Germany GPS	-0.03	0.52	Inf	-0.06	1.000
Canada GPS - Germany GPS	**-2.31**	**0.52**	**Inf**	**-4.48**	**< 0.001**

Supplementary Table 4.7: Tukey multiple comparisons of microbial diversity, Shannon diversity, Phylogenetic diversity and Gini-Simpson diversity between habitat of exposure and sample type Significant results are highlighted in bold, Df denotes degrees of freedom.

	Df	Shannon Diversity				Phylogenetic Diversity				Gini-Simpson Diversity			
		estimate	SE	t.ratio	p.value	estimate	SE	t.ratio	p.value	estimate	SE	t.ratio	p.value
Habitat of exposure: GPS													
Fish - Invertebrate	433	**-1.63**	**0.41**	**-3.93**	**< 0.001**	**-6.53**	**1.49**	**-4.39**	**< 0.001**	**-0.34**	**0.10**	**-3.52**	**< 0.01**
Fish - Water	432	0.07	0.55	0.12	0.992	1.23	1.99	0.62	0.812	-0.07	0.13	-0.54	0.853
Invertebrate - Water	431	**1.69**	**0.64**	**2.66**	**< 0.05**	**7.75**	**2.29**	**3.38**	**< 0.01**	0.27	0.15	1.82	0.163
Habitat of exposure: MAU													
Fish - Invertebrate	433	**-1.18**	**0.26**	**-4.55**	**< 0.001**	**-4.30**	**0.93**	**-4.61**	**< 0.001**	**-0.22**	**0.06**	**-3.59**	**< 0.01**
Fish - Water	432	-0.79	0.54	-1.45	0.314	2.13	1.95	1.09	0.518	-0.23	0.13	-1.86	0.151
Invertebrate - Water	431	0.39	0.56	0.70	0.765	**6.43**	**2.02**	**3.18**	**< 0.01**	-0.02	0.13	-0.13	0.991
Habitat of exposure: LAB													
Fish - Invertebrate	431	0.84	0.40	2.10	0.091	2.46	1.44	1.72	0.200	**0.23**	**0.09**	**2.50**	**< 0.05**
Fish - Water	431	**1.72**	**0.54**	**3.17**	**<0.01**	2.88	1.95	1.48	0.303	**0.43**	**0.13**	**3.41**	**< 0.01**
Invertebrate - Water	431	0.88	0.64	1.39	0.349	0.42	2.29	0.18	0.982	0.20	0.15	1.34	0.376
Sample Type: Fish													
GPS - MAU	432	**-0.49**	**0.14**	**-3.53**	**< 0.05**	**-1.46**	**0.50**	**-2.93**	**< 0.05**	**-0.09**	**0.03**	**-2.92**	**< 0.01**
GPS - LAB	428	**-1.23**	**0.24**	**-5.05**	**< 0.001**	**-4.14**	**0.88**	**-4.74**	**< 0.001**	**-0.26**	**0.06**	**-4.61**	**< 0.001**
MAU - LAB	431	**-0.74**	**0.21**	**-3.45**	**< 0.01**	**-2.69**	**0.77**	**-3.49**	**< 0.01**	**-0.17**	**0.05**	**-3.37**	**< 0.01**
Sample Type: Invertebrate													
GPS - MAU	431	-0.04	0.43	-0.10	0.995	0.77	1.53	0.50	0.869	0.03	0.10	0.28	0.958
GPS - LAB	431	**1.24**	**0.52**	**2.37**	**< 0.05**	**4.85**	**1.87**	**2.59**	**< 0.05**	**0.31**	**0.12**	**2.55**	**< 0.05**
MAU - LAB	431	**1.28**	**0.43**	**3.01**	**< 0.01**	**4.08**	**1.53**	**2.67**	**< 0.05**	**0.28**	**0.10**	**2.84**	**< 0.05**
Sample Type: Water													
GPS - MAU	431	-1.34	0.74	-1.83	0.163	-0.55	2.65	-0.21	0.976	-0.26	0.17	-1.52	0.284
GPS - LAB	431	0.43	0.74	0.58	0.831	-2.49	2.65	-0.94	0.616	0.23	0.17	1.38	0.352
MAU - LAB	431	**1.77**	**0.74**	**2.40**	**< 0.05**	-1.93	2.65	-0.73	0.746	**0.49**	**0.17**	**2.90**	**< 0.05**

271

Supplementary Table 4.8: Results from PERMANOVAs and betadisper for all levels of analysis. Month of collection used as block. Significant results are highlighted in bold, Df denotes degrees of freedom

Dataset	Factor	Df	Unweighted UniFrac Distance			Weighted UniFrac Distance			Bray-Curtis Dissimilarity		
			F-value	R2	P-value	F-value	R2	P-value	F-value	R2	P-value
All Sample Types - Block: Month of collection	**Sample Type**	**2**	**18.71**	**0.07**	**0.001**	**26.88**	**0.10**	**0.001**	**11.70**	**0.05**	**0.001**
	Habitat	**2**	**17.05**	**0.07**	**0.001**	**14.78**	**0.05**	**0.001**	**14.76**	**0.06**	**0.001**
	Sample Type x Habitat	**4**	**4.39**	**0.03**	**0.001**	**5.68**	**0.04**	**0.001**	**4.80**	**0.04**	**0.001**
	Residuals	433	-	0.83	-	-	0.80	-	-	0.86	-
	Total	441	-	1.00	-	-	1.00	-	-	1.00	-
	Dispersion Test	*Df*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*
	Sample Type dispersion	**2,439**	**20.14**	**999**	**< 0.001**	**7.80**	**999**	**< 0.001**	**59.29**	**999**	**< 0.001**
	Habitat dispersion	**2,439**	**8.70**	**999**	**< 0.001**	0.16	999	0.850	1.85	999	0.1588
	Sample Type x Habitat dispersion	**8,433**	**20.24**	**999**	**< 0.001**	**17.65**	**999**	**< 0.001**	**115.15**	**999**	**< 0.001**
Experimental mesocosms - Block: Month of collection	**Habitat**	**1**	**16.61**	**0.05**	**0.001**	**13.09**	**0.04**	**0.001**	**17.41**	**0.05**	**0.001**
	Country	**1**	**4.86**	**0.01**	**0.001**	**10.53**	**0.03**	**0.001**	**4.71**	**0.01**	**0.001**
	Ecotype	1	0.97	0.00	0.503	0.49	0.00	0.917	0.81	0.00	0.868
	I_{PI} group	**2**	**6.33**	**0.04**	**0.012**	**7.76**	**0.04**	**0.018**	**3.82**	**0.02**	**0.003**
	Habitat x Country	**1**	**4.04**	**0.01**	**0.003**	**4.91**	**0.01**	**0.002**	**3.90**	**0.01**	**0.001**
	Habitat x Ecotype	1	1.18	0.00	0.196	0.51	0.00	0.88	1.07	0.00	0.293
	Country x Ecotype	**1**	**1.57**	**0.00**	**0.044**	1.16	0.00	0.286	1.24	0.00	0.161
	Habitat x I_{PI} group	**2**	**1.85**	**0.01**	**0.009**	**2.18**	**0.01**	**0.03**	**1.52**	**0.01**	**0.022**
	Country x I_{PI} group	**2**	1.46	0.01	0.077	1.33	0.01	0.281	**1.37**	**0.01**	**0.041**
	Ecotype x I_{PI} group	2	1.32	0.01	0.085	0.81	0.00	0.7	1.06	0.01	0.364
	Habitat x Country x Ecotype	1	1.08	0.00	0.335	1.20	0.00	0.25	1.24	0.00	0.134
	Habitat x Country x I_{PI} group	1	1.01	0.00	0.406	1.26	0.00	0.246	1.02	0.00	0.384
	Habitat x Ecotype x I_{PI} group	1	0.96	0.00	0.418	0.61	0.00	0.801	1.00	0.00	0.438
	Country x Ecotype x I_{PI} group	2	0.91	0.01	0.547	1.04	0.01	0.371	1.07	0.01	0.235
	Residuals	300	-	0.84	-	-	0.83	-	-	0.86	-
	Total	319	-	1.00	-	-	1.00	-	-	1.00	-
	Dispersion Test	*Df*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*
	Habitat dispersion	**1,318**	**16.04**	**999**	**< 0.001**	0.35	999	0.554	3.30	999	0.070
	Country of Origin dispersion	**1,318**	**9.19**	**999**	**0.003**	**13.56**	**999**	**< 0.001**	**19.37**	**999**	**< 0.001**
	I_{PI} group dispersion	**2,317**	**3.96**	**999**	**0.020**	**19.63**	**999**	**< 0.001**	**8.12**	**999**	**< 0.001**

	Df									
Habitat x Country of Origin dispersion	**3,316**	**8.26**	**999**	**< 0.001**	**7.65**	**999**	**< 0.001**	**12.83**	**999**	**< 0.001**
Country x Ecotype dispersion	**3,316**	**3.32**	**999**	**0.020**	**6.17**	**999**	**< 0.001**	**6.25**	**999**	**< 0.001**
Habitat x I_{PI} group dispersion	**5,314**	**2.43**	**999**	**0.035**	**7.35**	**999**	**< 0.001**	**2.37**	**999**	**0.039**
Country of Origin x I_{PI} group dispersion	**5,314**	**4.24**	**999**	**< 0.001**	**13.65**	**999**	**< 0.001**	**9.75**	**999**	**< 0.001**
Ecotype x I_{PI} group dispersion	5,314	1.83	999	0.106	**8.11**	**999**	**< 0.001**	**4.33**	**999**	**< 0.001**

Großer Plöner See - Block: Month of collection

	Df	F-value	R2	Pr(>F)	F-value	R2	Pr(>F)	F-value	R2	Pr(>F)
Country	1	1.34	0.01	0.230	**3.74**	**0.03**	**0.014**	**2.61**	**0.02**	**0.010**
Ecotype	1	0.83	0.01	0.833	0.47	0.00	0.935	0.89	0.01	0.858
I_{PI} group GPS specific	2	1.41	0.02	0.556	0.91	0.02	0.891	1.48	0.03	0.810
Country x Ecotype	1	1.46	0.01	0.093	1.48	0.01	0.180	1.67	0.01	0.078
Country x I_{PI} group GPS specific	2	0.83	0.01	0.759	0.80	0.01	0.626	0.82	0.01	0.806
Ecotype x I_{PI} group GPS specific	2	0.89	0.02	0.572	0.73	0.01	0.657	0.97	0.02	0.389
Country x Ecotype x I_{PI} group GPS specific	2	0.78	0.01	0.834	0.49	0.01	0.926	0.77	0.01	0.856
Residuals	105	-	0.90	-	-	0.90	-	-	0.89	-
Total	116	-	1.00	-	-	1.00	-	-	1.00	-
Dispersion Test	*Df*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*
Country of Origin dispersion	1,115	-	-	-	0.18	999	0.6748	1.88	999	0.173

Malenter Au - Block: Month of collection

	Df	F-value	R2	Pr(>F)	F-value	R2	Pr(>F)	F-value	R2	Pr(>F)
Country	**1**	**7.88**	**0.04**	**0.001**	**12.16**	**0.05**	**0.001**	**6.26**	**0.03**	**0.001**
Ecotype	1	1.18	0.01	0.159	0.44	0.00	0.903	0.89	0.00	0.566
I_{PI} group MAU specific	**2**	**7.23**	**0.06**	**0.003**	**9.72**	**0.08**	**0.003**	**4.27**	**0.04**	**0.003**
Country x Ecotype	1	1.17	0.01	0.151	1.10	0.00	0.22	1.11	0.01	0.203
Country x I_{PI} group MAU specific	2	1.38	0.01	0.188	1.51	0.01	0.24	**1.51**	**0.01**	**0.017**
Ecotype x I_{PI} group MAU specific	2	0.92	0.01	0.560	0.86	0.01	0.527	1.12	0.01	0.187
Country x Ecotype x I_{PI} group MAU specific	2	1.81	0.02	0.073	1.81	0.02	0.077	1.34	0.01	0.087
Residuals	191	-	0.85	-	-	0.82	-	-	0.89	-
Total	202	-	1.00	-	-	1.00	-	-	1.00	-
Dispersion Test	*Df*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*	*F-value*	*N.Perm*	*Pr(>F)*
Country of Origin dispersion	**1,201**	**23.10**	**999**	**< 0.001**	**19.74**	**999**	**< 0.001**	**30.36**	**999**	**< 0.001**
I_{PI} group MAU specific dispersion	**2,200**	**4.14**	**999**	**0.017**	**20.39**	**999**	**< 0.001**	**4.51**	**999**	**0.012**
Country x I_{PI} group MAU specific dispersion	**5,197**	-	-	-	-	-	-	**2.82**	**999**	**0.017**
Country x Ecotype x I_{PI} group MAU specific dispersion	**11,191**	**4.72**	**999**	**< 0.001**	-	-	-	**5.72**	**999**	**< 0.001**

Supplementary Table 4.9: Pairwise PERMANOVA results from pairwise.adonis2 function for paired habitat of exposure and sample type comparisons. Month of collection used as block. Significant results are highlighted in bold, Df denotes degrees of freedom

Factors	Df	Unweighted UniFrac Distance					Weighted UniFrac Distance					Bray-Curtis Dissimilarity				
		Sums Of Sqs	Mean Sqs	F.Model	R2	Pr(>F)	Sums Of Sqs	Mean Sqs	F.Model	R2	Pr(>F)	Sums Of Sqs	Mean Sqs	F.Model	R2	Pr(>F)
Habitat of exposure: GPS																
Fish - Invertebrate	1,126	1.90	1.90	7.70	0.06	**0.001**	2.73	2.73	11.16	0.08	**0.001**	2.35	2.35	5.94	0.05	**0.001**
Fish - Water	1,121	2.25	2.25	9.26	0.07	**0.001**	2.76	2.76	11.30	0.09	**0.001**	2.83	2.83	7.34	0.06	**0.002**
Invertebrate - Water	1,14	1.69	1.69	10.49	0.45	**0.001**	2.26	2.26	21.07	0.62	**0.001**	2.17	2.17	8.59	0.40	**0.001**
Habitat of exposure: MAU																
Fish - Invertebrate	1,232	4.86	4.86	17.57	0.07	**0.001**	6.54	6.54	25.41	0.10	**0.001**	4.58	4.58	11.16	0.05	**0.001**
Fish - Water	1,207	2.01	2.01	7.21	0.03	**0.001**	2.01	2.01	7.56	0.04	**0.002**	2.47	2.47	6.10	0.03	**0.001**
Invertebrate - Water	1,34	1.78	1.78	8.43	0.20	**0.001**	1.28	1.28	8.65	0.21	**0.001**	2.33	2.33	6.65	0.17	**0.001**
Habitat of exposure: LAB																
Fish - Invertebrate	1,66	1.69	1.69	7.03	0.10	**0.001**	1.82	1.82	7.85	0.11	**0.001**	2.08	2.08	5.34	0.08	**0.001**
Fish - Water	1,61	1.72	1.72	7.65	0.11	**0.001**	3.15	3.15	14.67	0.20	**0.001**	2.54	2.54	6.82	0.10	**0.001**
Invertebrate - Water	1,14	1.16	1.16	4.79	0.27	**0.001**	2.16	2.16	12.11	0.48	**0.001**	2.05	2.05	7.62	0.37	**0.001**
Sample Type: Fish																
GPS - MAU	1,319	4.22	4.22	15.57	0.05	**0.001**	3.17	3.17	12.00	0.04	**0.001**	6.78	6.78	16.64	0.05	**0.001**
GPS - LAB	1,173	4.75	4.75	19.50	0.10	**0.001**	5.05	5.05	20.67	0.11	**0.001**	5.75	5.75	14.49	0.08	**0.001**
MAU - LAB	1,259	5.07	5.07	18.65	0.07	**0.001**	3.89	3.89	14.87	0.05	**0.001**	6.09	6.09	14.89	0.05	**0.001**
Sample Type: Invertebrate																
GPS - MAU	1,39	0.90	0.90	4.01	0.10	**0.001**	0.32	0.32	1.97	0.05	**0.036**	1.22	1.22	3.19	0.08	**0.001**
GPS - LAB	1,19	1.14	1.13	4.56	0.20	**0.001**	1.32	1.32	6.54	0.27	**0.001**	1.36	1.36	3.79	0.17	**0.001**
MAU - LAB	1,39	1.38	1.38	5.66	0.13	**0.001**	1.77	1.77	9.53	0.20	**0.001**	1.74	1.74	4.53	0.11	**0.001**
Sample Type: Water																
GPS - MAU	1,9	0.43	0.43	6.51	0.45	**0.006**	0.53	0.53	40.50	0.84	**0.006**	1.06	1.06	28.66	0.78	**0.013**
GPS - LAB	1,9	1.17	1.17	12.43	0.61	**0.009**	2.20	2.20	215.48	0.96	**0.011**	2.23	2.23	52.13	0.87	**0.015**
MAU - LAB	1,9	0.99	0.99	9.73	0.55	**0.006**	1.50	1.50	80.96	0.91	**0.009**	2.13	2.13	35.17	0.81	**0.008**

Supplementary Table 4.10: Pairwise PERMANOVA results from pairwise.adonis2 function for paired habitat of exposure and I_{PI} group comparisons. Month of collection used as block. Significant results are highlighted in bold, Df denotes degrees of freedom.

Comparison	Df	Unweighted UniFrac			Weighted UniFrac			Bray-Curtis		
		F.Model	R2	Pr(>F)	F.Model	R2	Pr(>F)	F.Model	R2	Pr(>F)
GPS Low vs GPS High	1,50	0.61	0.01	0.960	0.16	0.00	0.999	0.57	0.01	0.984
GPS Medium vs GPS High	1,109	1.89	0.02	0.502	1.49	0.01	0.670	2.11	0.02	0.698
GPS Medium vs GPS Low	1,69	1.12	0.02	0.273	0.78	0.01	0.449	0.99	0.01	0.366
GPS Low vs MAU High	1,64	1.70	0.03	0.232	1.70	0.03	0.333	1.39	0.02	0.069
GPS Medium vs MAU High	**1,123**	**12.95**	**0.10**	**0.001**	**14.32**	**0.10**	**0.001**	**9.70**	**0.07**	**0.001**
MAU High vs GPS High	**1,104**	**7.05**	**0.06**	**0.001**	**7.56**	**0.07**	**0.001**	**4.67**	**0.04**	**0.001**
MAU Low vs GPS High	**1,145**	**9.31**	**0.06**	**0.001**	**7.97**	**0.05**	**0.001**	**8.90**	**0.06**	**0.001**
MAU Low vs GPS Low	**1,105**	**2.08**	**0.02**	**0.004**	**2.02**	**0.02**	**0.033**	**2.36**	**0.02**	**0.001**
MAU Low vs GPS Medium	**1,164**	**12.60**	**0.07**	**0.001**	**9.49**	**0.05**	**0.001**	**13.99**	**0.08**	**0.001**
MAU Medium vs GPS High	**1,86**	**5.60**	**0.06**	**0.001**	**3.68**	**0.04**	**0.002**	**5.14**	**0.06**	**0.001**
MAU Medium vs GPS Low	**1,46**	**1.78**	**0.04**	**0.027**	1.15	0.02	0.238	**1.86**	**0.04**	**0.002**
MAU Medium vs GPS Medium	**1,105**	**7.04**	**0.06**	**0.001**	**5.21**	**0.05**	**0.002**	**8.03**	**0.07**	**0.001**
MAU Low vs MAU High	**1,159**	12.63	0.07	0.053	17.55	0.10	0.060	**6.97**	**0.04**	**0.005**
MAU Medium vs MAU High	**1,100**	**6.09**	**0.06**	**0.007**	**7.53**	**0.07**	**0.009**	**3.95**	**0.04**	**0.001**
MAU Medium vs MAU Low	1,141	1.36	0.01	0.814	0.97	0.01	0.969	1.27	0.01	0.908

Supplementary Table 4.11: Pairwise PERMANOVA results from pairwise.adonis2 function for paired habitat of exposure and country of origin comparisons. Month of collection used as block. Significant results are highlighted in bold, Df denotes degrees of freedom.

Pairwise comparisons	Df	Unweighted UniFrac Distance					Weighted UniFrac Distance					Bray-Curtis Dissimilarity				
		Sums Of Sqs	Mean Sqs	F value	R2	Pr(>F)	Sums Of Sqs	Mean Sqs	F value	R2	Pr(>F)	Sums Of Sqs	Mean Sqs	F value	R2	Pr(>F)
Habitat of Exposure																
MAU																
German - Canadian	**1,202**	**2.01**	**2.01**	**7.33**	**0.04**	**0.001**	**2.86**	**2.86**	**11.10**	**0.05**	**0.001**	**2.42**	**2.42**	**6.01**	**0.03**	**0.001**
GPS																
German - Canadian	**1,116**	0.33	0.33	1.34	0.01	0.231	**0.94**	**0.94**	**3.82**	**0.03**	**0.015**	**1.02**	**1.02**	**2.60**	**0.02**	**0.007**
Country of Origin																
German																
MAU - GPS	**1,143**	**2.72**	**2.72**	**10.92**	**0.07**	**0.001**	**1.81**	**1.81**	**8.01**	**0.05**	**0.001**	**4.15**	**4.15**	**10.84**	**0.07**	**0.001**
Canadian																
MAU - GPS	**1,175**	**2.54**	**2.54**	**9.16**	**0.05**	**0.001**	**2.54**	**2.54**	**9.22**	**0.05**	**0.001**	**4.34**	**4.33**	**10.50**	**0.06**	**0.001**

Supplementary Table 4.12: Significant SIMPER results from habitat of exposure (MAU – Malenter Au, GPS – Großer Plöner See) and country of origin comparisons (GER = Germany, CAN = Canada).

Comparison (X_Y)	ASV	SIMPER	Pr(>F)	Mean Abundance X ± SD	Mean Abundance Y ± SD	Phylum	Class	Order	Family	Genus	Species
MAU Can_GPS Can	bee63189d08003ab4d401cd1d3aeab91	0.04	0.002	0.062 ± 0.213	0.000 ± 0.001	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	*NA*
MAU Can_GPS Ger	bee63189d08003ab4d401cd1d3aeab91	0.05	0.003	0.062 ± 0.213	0.013 ± 0.097	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	*NA*
MAU Ger_GPS Ger	bee63189d08003ab4d401cd1d3aeab91	0.01	0.006	0.010 ± 0.069	0.013 ± 0.097	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	*NA*
MAU Can_GPS Can	f3f475d3bd572281d12d2fe3783df534	0.05	0.001	0.087 ± 0.213	0.001 ± 0.002	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	*Carnobacterium*	*viridans*
MAU Ger_MAU Can	cfa1eddaf5f286cf0a8d9742017dad8e	0.01	0.010	0.017 ± 0.035	0.012 ± 0.025	Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	*NA*	*NA*
GPS Can_GPS Ger	a2c2577cd7b15660d4d0d421a4213bd3	0.01	0.036	0.019 ± 0.027	0.009 ± 0.018	Proteobacteria	Alphaproteobacteria	Rhizobiales	*NA*	*NA*	*NA*
MAU Ger_GPS Ger	0938f5e4e046f002176f9baabcbe2491	0.01	< 0.001	0.000 ± 0.001	0.019 ± 0.027	Cyanobacteria	Synechococcophycideae	Synechococcales	Synechococcaceae	*Synechococcus*	*NA*
MAU Ger_MAU Can	f3f475d3bd572281d12d2fe3783df534	0.07	< 0.001	0.004 ± 0.032	0.087 ± 0.213	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	*Carnobacterium*	*viridans*
MAU Can_GPS Ger	f3f475d3bd572281d12d2fe3783df534	0.06	< 0.001	0.087 ± 0.213	0.006 ± 0.026	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	*Carnobacterium*	*viridans*
MAU Ger_GPS Ger	cfa1eddaf5f286cf0a8d9742017dad8e	0.01	< 0.001	0.017 ± 0.035	0.002 ± 0.010	Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	*NA*	*NA*
MAU Ger_GPS Ger	3f5225a8b2e87a448aab5f44a4412c2a	0.01	< 0.001	0.000 ± 0.000	0.021 ± 0.037	Planctomycetes	Planctomycetia	Gemmatales	Isosphaeraceae	*NA*	*NA*
MAU Can_GPS Can	a2c2577cd7b15660d4d0d421a4213bd3	0.01	< 0.001	0.000 ± 0.000	0.019 ± 0.027	Proteobacteria	Alphaproteobacteria	Rhizobiales	*NA*	*NA*	*NA*
MAU Ger_GPS Can	a2c2577cd7b15660d4d0d421a4213bd3	0.01	< 0.001	0.000 ± 0.000	0.019 ± 0.027	Proteobacteria	Alphaproteobacteria	Rhizobiales	*NA*	*NA*	*NA*
MAU Ger_MAU Can	32fa112a32f7293d2b0c5bf9c375ab80	0.03	< 0.001	0.001 ± 0.003	0.045 ± 0.150	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	*Phyllobacterium*	*NA*
MAU Can_GPS Can	32fa112a32f7293d2b0c5bf9c375ab80	0.02	< 0.001	0.045 ± 0.150	0.000 ± 0.001	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	*Phyllobacterium*	*NA*
MAU Can_GPS Ger	32fa112a32f7293d2b0c5bf9c375ab80	0.03	< 0.001	0.045 ± 0.150	0.006 ± 0.040	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	*Phyllobacterium*	*NA*
MAU Can_GPS Ger	2441d7d478b956c8f710f23327981381	0.06	< 0.001	0.002 ± 0.008	0.114 ± 0.203	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
MAU Ger_GPS Ger	2441d7d478b956c8f710f23327981381	0.07	< 0.001	0.002 ± 0.007	0.114 ± 0.203	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
MAU Ger_MAU Can	284dd9dfc7026850bffe4b81c9b284cc	0.06	0.011	0.000 ± 0.002	0.067 ± 0.210	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Serratia*	*NA*
MAU Can_GPS Can	284dd9dfc7026850bffe4b81c9b284cc	0.05	0.039	0.067 ± 0.210	0.000 ± 0.002	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Serratia*	*NA*
MAU Ger_MAU Can	24fb88dd51e0a967e75c3a33a1a86d47	0.01	< 0.001	0.015 ± 0.031	0.005 ± 0.034	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
MAU Ger_MAU Can	cab81200dc8a1b7011cb421b1df82262	0.03	0.004	0.037 ± 0.041	0.033 ± 0.071	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
MAU Ger_MAU Can	85a8b29dd3c64524083b9582244855b7	0.01	< 0.001	0.025 ± 0.028	0.008 ± 0.021	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
MAU Can_GPS Ger	85a8b29dd3c64524083b9582244855b7	0.02	< 0.001	0.008 ± 0.021	0.034 ± 0.060	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
MAU Can_GPS Can	85a8b29dd3c64524083b9582244855b7	0.01	< 0.001	0.008 ± 0.021	0.030 ± 0.057	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
MAU Can_GPS Can	47a55f7ba97286602de71a9eab0f5c1b	0.25	< 0.001	0.019 ± 0.090	0.324 ± 0.421	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
MAU Ger_GPS Can	47a55f7ba97286602de71a9eab0f5c1b	0.27	< 0.001	0.027 ± 0.089	0.324 ± 0.421	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
MAU Can_GPS Ger	47a55f7ba97286602de71a9eab0f5c1b	0.13	< 0.001	0.019 ± 0.090	0.157 ± 0.330	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
GPS Can_GPS Ger	47a55f7ba97286602de71a9eab0f5c1b	0.37	< 0.001	0.324 ± 0.421	0.157 ± 0.330	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
MAU Ger_GPS Ger	47a55f7ba97286602de71a9eab0f5c1b	0.15	< 0.001	0.027 ± 0.089	0.157 ± 0.330	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
MAU Ger_GPS Ger	cab81200dc8a1b7011cb421b1df82262	0.02	< 0.001	0.037 ± 0.041	0.003 ± 0.016	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*

Comparison (X_Y)	ASV	SIMPER	Pr(>F)	Mean Abundance X ± SD	Mean Abundance Y ± SD	Phylum	Class	Order	Family	Genus	Species
MAU Ger_GPS Can	cab81200dc8a1b7011cb421b1df82262	0.02	< 0.001	0.037 ± 0.041	0.002 ± 0.015	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
MAU Can_GPS Can	4fb4be24dfbc56d883e1ec8f222565b3	0.02	0.023	0.016 ± 0.060	0.018 ± 0.127	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	*Acinetobacter*	*lwoffii*
MAU Can_GPS Ger	cab81200dc8a1b7011cb421b1df82262	0.02	< 0.001	0.033 ± 0.071	0.003 ± 0.016	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
MAU Can_GPS Can	cab81200dc8a1b7011cb421b1df82262	0.01	< 0.001	0.033 ± 0.071	0.002 ± 0.015	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
MAU Ger_MAU Can	97430edb43f5f48e4b7a8ae996f4e799	0.01	0.002	0.015 ± 0.053	0.006 ± 0.042	Tenericutes	CK-1C4-19	*NA*	*NA*	*NA*	*NA*
MAU Ger_GPS Ger	97430edb43f5f48e4b7a8ae996f4e799	0.01	< 0.001	0.015 ± 0.053	0.001 ± 0.006	Tenericutes	CK-1C4-19	*NA*	*NA*	*NA*	*NA*
MAU Ger_MAU Can	4fd952db1056e1588390cc62e351a2cc	0.01	< 0.001	0.022 ± 0.024	0.009 ± 0.017	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	*Luteolibacter*	*NA*
MAU Ger_GPS Ger	4fd952db1056e1588390cc62e351a2cc	0.01	< 0.001	0.022 ± 0.024	0.010 ± 0.019	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	*Luteolibacter*	*NA*
MAU Ger_GPS Can	4fd952db1056e1588390cc62e351a2cc	0.01	< 0.001	0.022 ± 0.024	0.011 ± 0.024	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	*Luteolibacter*	*NA*

Supplementary Table 4.13: Significant SIMPER results for Malenter Au only dataset between country of origin (GER = Germany, CAN = Canada) and I$_{PI}$ group (L = Low, M = Medium, H = High) pairs.

| Comparison (X_Y)	ASV	SIMPER	Pr(>F)	Mean Abundance X ± SD	Mean Abundance Y ± SD	Phylum	Class	Order	Family	Genus	Species
GER M_CAN H	24fb88dd51e0a967e75c3a33a1a86d47	0.01	0.019	0.018 ± 0.043	0.001 ± 0.004	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
CAN L_GER L	24fb88dd51e0a967e75c3a33a1a86d47	0.01	<0.001	0.008 ± 0.046	0.021 ± 0.028	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER L_CAN H	24fb88dd51e0a967e75c3a33a1a86d47	0.01	<0.001	0.021 ± 0.028	0.001 ± 0.004	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER L_GER H	24fb88dd51e0a967e75c3a33a1a86d47	0.01	0.001	0.021 ± 0.028	0.000 ± 0.001	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER L_CAN M	24fb88dd51e0a967e75c3a33a1a86d47	0.01	0.006	0.021 ± 0.028	0.007 ± 0.031	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER M_CAN M	284dd9dfc7026850bffe4b81c9b284cc	0.13	0.043	0.000 ± 0.000	0.137 ± 0.301	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Serratia*	*NA*
GER L_CAN M	284dd9dfc7026850bffe4b81c9b284cc	0.12	0.029	0.000 ± 0.002	0.137 ± 0.301	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Serratia*	*NA*
GER M_GER L	32b3ec26b5e462171abfd24998eae42f	0.01	0.033	0.011 ± 0.014	0.024 ± 0.021	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
CAN L_GER L	32b3ec26b5e462171abfd24998eae42f	0.01	<0.001	0.005 ± 0.010	0.024 ± 0.021	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER L_CAN H	32b3ec26b5e462171abfd24998eae42f	0.01	<0.001	0.024 ± 0.021	0.002 ± 0.006	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER L_GER H	32b3ec26b5e462171abfd24998eae42f	0.01	0.001	0.024 ± 0.021	0.010 ± 0.024	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER M_CAN M	32fa112a32f7293d2b0c5bf9c375ab80	0.04	0.025	0.000 ± 0.001	0.056 ± 0.200	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	*Phyllobacterium*	*NA*
CAN L_GER L	32fa112a32f7293d2b0c5bf9c375ab80	0.04	0.004	0.062 ± 0.173	0.000 ± 0.002	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	*Phyllobacterium*	*NA*
GER L_CAN M	32fa112a32f7293d2b0c5bf9c375ab80	0.04	<0.001	0.000 ± 0.002	0.056 ± 0.200	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	*Phyllobacterium*	*NA*
GER H_CAN M	32fa112a32f7293d2b0c5bf9c375ab80	0.04	0.036	0.002 ± 0.006	0.056 ± 0.200	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	*Phyllobacterium*	*NA*
GER M_CAN L	4fd952db1056e1588390cc62e351a2cc	0.01	0.027	0.022 ± 0.020	0.009 ± 0.018	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	*Luteolibacter*	*NA*
GER M_CAN H	4fd952db1056e1588390cc62e351a2cc	0.01	<0.001	0.022 ± 0.020	0.006 ± 0.016	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	*Luteolibacter*	*NA*
GER M_GER H	4fd952db1056e1588390cc62e351a2cc	0.02	0.001	0.022 ± 0.020	0.003 ± 0.007	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	*Luteolibacter*	*NA*
CAN L_GER L	4fd952db1056e1588390cc62e351a2cc	0.02	<0.001	0.009 ± 0.018	0.030 ± 0.025	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	*Luteolibacter*	*NA*

						Phylum	Class	Order	Family	Genus	Species
GER L_CAN H	4fd952db1056e1588390cc62e351a2cc	0.02	<0.001	0.030 ± 0.025	0.006 ± 0.016	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	*Luteolibacter*	*NA*
GER L_GER H	4fd952db1056e1588390cc62e351a2cc	0.02	<0.001	0.030 ± 0.025	0.003 ± 0.007	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	*Luteolibacter*	*NA*
GER M_CAN L	85a8b29dd3c64524083b9582244855b7	0.02	<0.001	0.033 ± 0.033	0.007 ± 0.015	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER M_CAN H	85a8b29dd3c64524083b9582244855b7	0.02	<0.001	0.033 ± 0.033	0.007 ± 0.028	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER M_GER H	85a8b29dd3c64524083b9582244855b7	0.02	0.005	0.033 ± 0.033	0.011 ± 0.033	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
CAN L_GER L	85a8b29dd3c64524083b9582244855b7	0.01	<0.001	0.007 ± 0.015	0.028 ± 0.020	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER L_CAN H	85a8b29dd3c64524083b9582244855b7	0.02	<0.001	0.028 ± 0.020	0.007 ± 0.028	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER L_GER H	85a8b29dd3c64524083b9582244855b7	0.02	<0.001	0.028 ± 0.020	0.011 ± 0.033	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER L_CAN M	85a8b29dd3c64524083b9582244855b7	0.01	0.039	0.028 ± 0.020	0.012 ± 0.017	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER M_GER H	8924b708d46794fceeadb26d7a702125	0.01	0.003	0.018 ± 0.016	0.007 ± 0.013	Proteobacteria	Alphaproteobacteria	Rhizobiales	*NA*	*NA*	*NA*
GER L_GER H	8924b708d46794fceeadb26d7a702125	0.01	<0.001	0.018 ± 0.014	0.007 ± 0.013	Proteobacteria	Alphaproteobacteria	Rhizobiales	*NA*	*NA*	*NA*
GER M_CAN H	97430edb43f5f48e4b7a8ae996f4e799	0.01	<0.001	0.016 ± 0.035	0.000 ± 0.001	Tenericutes	CK-1C4-19	*NA*	*NA*	*NA*	*NA*
GER M_GER H	97430edb43f5f48e4b7a8ae996f4e799	0.01	0.039	0.016 ± 0.035	0.001 ± 0.003	Tenericutes	CK-1C4-19	*NA*	*NA*	*NA*	*NA*
CAN L_GER L	97430edb43f5f48e4b7a8ae996f4e799	0.02	0.026	0.012 ± 0.061	0.020 ± 0.068	Tenericutes	CK-1C4-19	*NA*	*NA*	*NA*	*NA*
GER L_CAN H	97430edb43f5f48e4b7a8ae996f4e799	0.02	<0.001	0.020 ± 0.068	0.000 ± 0.001	Tenericutes	CK-1C4-19	*NA*	*NA*	*NA*	*NA*
GER L_GER H	97430edb43f5f48e4b7a8ae996f4e799	0.02	0.006	0.020 ± 0.068	0.001 ± 0.003	Tenericutes	CK-1C4-19	*NA*	*NA*	*NA*	*NA*
GER M_GER L	a9c54df0ff16633d123cd932ccfb6155	0.01	0.029	0.003 ± 0.007	0.014 ± 0.020	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
GER L_CAN H	a9c54df0ff16633d123cd932ccfb6155	0.01	<0.001	0.014 ± 0.020	0.005 ± 0.023	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
GER L_GER H	a9c54df0ff16633d123cd932ccfb6155	0.01	0.036	0.014 ± 0.020	0.004 ± 0.009	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
GER L_CAN H	b9bf39824afdab3ae1a49ae1d416d551	0.01	0.009	0.020 ± 0.057	0.000 ± 0.000	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	*NA*	*NA*
CAN L_CAN M	bee63189d08003ab4d401cd1d3aeab91	0.12	0.045	0.033 ± 0.163	0.125 ± 0.314	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	*NA*
GER M_GER L	cab81200dc8a1b7011cb421b1df82262	0.03	0.008	0.023 ± 0.031	0.053 ± 0.045	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
CAN L_GER L	cab81200dc8a1b7011cb421b1df82262	0.03	<0.001	0.038 ± 0.066	0.053 ± 0.045	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
CAN L_CAN H	cab81200dc8a1b7011cb421b1df82262	0.03	0.025	0.038 ± 0.066	0.031 ± 0.090	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
GER L_CAN H	cab81200dc8a1b7011cb421b1df82262	0.03	<0.001	0.053 ± 0.045	0.031 ± 0.090	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
GER L_GER H	cab81200dc8a1b7011cb421b1df82262	0.03	<0.001	0.053 ± 0.045	0.016 ± 0.027	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	*Rickettsiella*	*NA*
CAN L_CAN H	cfa1eddaf5f286cf0a8d9742017dad8e	0.01	<0.001	0.021 ± 0.030	0.002 ± 0.009	Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	*NA*	*NA*
CAN L_GER H	cfa1eddaf5f286cf0a8d9742017dad8e	0.01	0.003	0.021 ± 0.030	0.001 ± 0.004	Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	*NA*	*NA*
GER L_CAN H	cfa1eddaf5f286cf0a8d9742017dad8e	0.02	<0.001	0.026 ± 0.045	0.002 ± 0.009	Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	*NA*	*NA*
GER L_GER H	cfa1eddaf5f286cf0a8d9742017dad8e	0.02	<0.001	0.026 ± 0.045	0.001 ± 0.004	Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	*NA*	*NA*
GER L_CAN M	cfa1eddaf5f286cf0a8d9742017dad8e	0.02	0.021	0.026 ± 0.045	0.007 ± 0.021	Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	*NA*	*NA*
GER M_CAN H	dd9db68f84bf9527eb34886a795e08e5	0.01	0.027	0.015 ± 0.043	0.003 ± 0.010	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
GER M_GER H	dd9db68f84bf9527eb34886a795e08e5	0.01	0.027	0.015 ± 0.043	0.000 ± 0.000	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Rhodobacter*	*NA*
CAN L_CAN H	e641108719ce19f0f12d52925319d8d6	0.06	0.004	0.000 ± 0.000	0.098 ± 0.276	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*NA*	*NA*
CAN L_CAN M	e641108719ce19f0f12d52925319d8d6	0.05	0.020	0.000 ± 0.000	0.058 ± 0.193	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*NA*	*NA*
GER M_CAN L	f3f475d3bd572281d12d2fe3783df534	0.08	0.020	0.000 ± 0.001	0.099 ± 0.210	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	*Carnobacterium*	*viridans*
GER M_CAN H	f3f475d3bd572281d12d2fe3783df534	0.07	0.036	0.000 ± 0.001	0.108 ± 0.263	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	*Carnobacterium*	*viridans*

GER M_CAN M	f3f475d3bd572281d12d2fe3783df534	0.01	0.016	0.000 ± 0.001	0.015 ± 0.030	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	*Carnobacterium*	*viridans*
CAN L_GER L	f3f475d3bd572281d12d2fe3783df534	0.08	0.007	0.099 ± 0.210	0.001 ± 0.003	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	*Carnobacterium*	*viridans*
CAN L_GER H	f3f475d3bd572281d12d2fe3783df534	0.09	0.033	0.099 ± 0.210	0.016 ± 0.068	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	*Carnobacterium*	*viridans*
GER L_CAN H	f3f475d3bd572281d12d2fe3783df534	0.06	0.027	0.001 ± 0.003	0.108 ± 0.263	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	*Carnobacterium*	*viridans*
GER L_CAN M	f3f475d3bd572281d12d2fe3783df534	0.01	0.009	0.001 ± 0.003	0.015 ± 0.030	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	*Carnobacterium*	*viridans*
GER H_CAN M	f3f475d3bd572281d12d2fe3783df534	0.02	0.036	0.014 ± 0.068	0.015 ± 0.030	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	*Carnobacterium*	*viridans*

Supplementary Materials – Chapter 5

Supplementary Table 5.1: Number of 16S rRNA samples: dry swabs, wet swabs, sand and extraction blanks removed prior to analysis as they contained fewer than 500 reads.

| | Island | | | | | |
Sample Type	Boa Vista	Maio	Sal	Santo Antão	NA	Total
Dry Swab	1	1	2	3	-	7
Wet Swab	5	8	10	8	-	31
Sand	-	-	0	0	-	0
Extraction Blank	-	-	-	-	6	6
Total	5	8	10	8	6	44
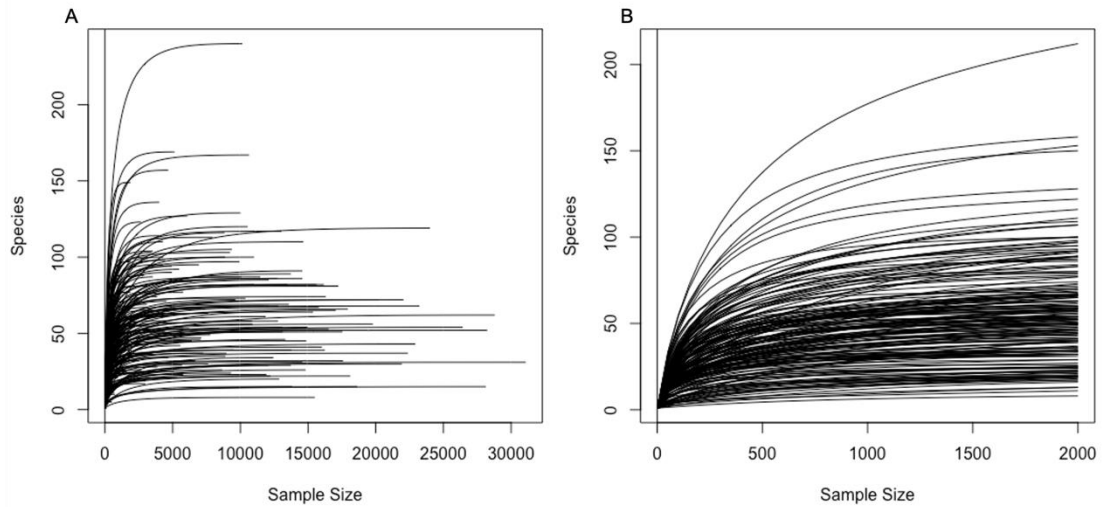
Supplementary Figure 5.1: A) Rarefaction curve for all microbial samples >500 reads. B) Rarefaction curve for microbial samples rarefied to 2000 reads

Supplementary Table 5.2: Results of Similarity Percentage Analysis (SIMPER) between dry and wet swab samples to identify ASVs driving differences between groups. Taxonomic information for each ASV included.

Comparison (X_Y)	ASV	p.value	Mean Abundance X ± SD	Mean Abundance Y ± SD	Phylum	Class	Order	Family	Genus	Species
Non-Rarefied										
Dry_Wet	95064b90639c6c20f85c141b5e8d6a15	0.001	0.013 ± 0.037	0.009 ± 0.025	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	*NA*	*NA*
Dry_Wet	95c95bb6bdcc327d75c06f8b3472723f	0.001	0.008 ± 0.052	0.046 ± 0.149	Proteobacteria	Alphaproteobacteria	Rhizobiales	*NA*	*NA*	*NA*
Dry_Wet	f62dc709536edba0135034e740915a92	< 0.001	0.018 ± 0.062	0.005 ± 0.026	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	*Helicobacter*	*NA*
Dry_Wet	ad7f3efd185f47aab4f7badf2c822cc8	< 0.001	0.014 ± 0.027	0.012 ± 0.043	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	*Helicobacter*	*NA*
Dry_Wet	0015124e5777b57174a1af7411703bbb	0.042	0.015 ± 0.035	0.012 ± 0.024	Proteobacteria	Gammaproteobacteria	Pasteurellales	Pasteurellaceae	*NA*	*NA*
Dry_Wet	67b0cf6af26b9fd959d860a52a65bc58	0.023	0.032 ± 0.089	0.028 ± 0.084	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Photobacterium*	*damselae*
Dry_Wet	9be79137deae3d1b144b96c452427950	0.004	0.016 ± 0.043	0.011 ± 0.018	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*	*NA*
Dry_Wet	6432d50fbe1bdecc92b8669a98cdcdcd	< 0.001	0.024 ± 0.055	0.015 ± 0.061	Spirochaetes	[Brachyspirae]	[Brachyspirales]	Brachyspiraceae	*Brachyspira*	*NA*
Rarefied										
Dry_Wet	95c95bb6bdcc327d75c06f8b3472723f	0.002	0.006 ± 0.054	0.046 ± 0.163	Proteobacteria	Alphaproteobacteria	Rhizobiales	*NA*	*NA*	*NA*
Dry_Wet	f62dc709536edba0135034e740915a92	0.003	0.018 ± 0.064	0.005 ± 0.033	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	*Helicobacter*	*NA*
Dry_Wet	6432d50fbe1bdecc92b8669a98cdcdcd	0.003	0.020 ± 0.047	0.013 ± 0.061	Spirochaetes	[Brachyspirae]	[Brachyspirales]	Brachyspiraceae	*Brachyspira*	*NA*

Supplementary Table 5.3: Pairwise post hoc Tukey tests between turtle nesting island and infection groups. Significant results in bold.

	Pairwise comparison	estimate	SE	Df	t.ratio	Adjusted p value
Non-rarefied Dry swabs	Boa Vista non-infected - Maio non-infected	-0.11	0.08	118	-1.34	0.880
	Boa Vista non-infected - Sal non-infected	-0.18	0.08	118	-2.42	0.241
	Boa Vista non-infected - Santo Antão non-infected	-0.22	0.08	118	-2.65	0.150
	Boa Vista non-infected - Boa Vista infected	-0.09	0.07	117	-1.26	0.911
	Boa Vista non-infected - Maio infected	-0.19	0.07	118	-2.59	0.169
	Boa Vista non-infected - Sal infected	-0.12	0.07	115	-1.74	0.659
	Boa Vista non-infected - Santo Antão infected	-0.10	0.07	118	-1.35	0.879
	Maio non-infected - Sal non-infected	-0.07	0.06	117	-1.20	0.930
	Maio non-infected - Santo Antão non-infected	-0.11	0.07	117	-1.59	0.756
	Maio non-infected - Boa Vista infected	0.02	0.06	116	0.28	1.000
	Maio non-infected - Maio infected	-0.08	0.06	116	-1.38	0.864
	Maio non-infected - Sal infected	-0.02	0.05	117	-0.28	1.000
	Maio non-infected - Santo Antão infected	0.01	0.06	117	0.14	1.000
	Sal non-infected - Santo Antão non-infected	-0.03	0.06	117	-0.57	0.999
	Sal non-infected - Boa Vista infected	0.09	0.05	118	1.80	0.624
	Sal non-infected - Maio infected	-0.01	0.05	118	-0.14	1.000
	Sal non-infected - Sal infected	0.06	0.05	118	1.26	0.911
	Sal non-infected - Santo Antão infected	0.08	0.05	117	1.54	0.785
	Santo Antão non-infected - Boa Vista infected	0.12	0.06	117	2.15	0.393
	Santo Antão non-infected - Maio infected	0.03	0.06	118	0.47	1.000
	Santo Antão non-infected - Sal infected	0.09	0.05	117	1.71	0.684
	Santo Antão non-infected - Santo Antão infected	0.12	0.06	118	1.91	0.548
	Boa Vista infected - Maio infected	-0.10	0.05	117	-2.09	0.429
	Boa Vista infected - Sal infected	-0.03	0.04	116	-0.78	0.994
	Boa Vista infected - Santo Antão infected	-0.01	0.05	118	-0.16	1.000
	Maio infected - Sal infected	0.07	0.04	118	1.55	0.780
	Maio infected - Santo Antão infected	0.09	0.05	117	1.82	0.608
	Sal infected - Santo Antão infected	0.02	0.04	118	0.53	1.000
Rarefied Dry swabs	Boa Vista non-infected - Maio non-infected	-0.07	0.52	102	-0.14	1.000
	Boa Vista non-infected - Sal non-infected	-0.69	0.48	102	-1.43	0.841
	Boa Vista non-infected - Santo Antão non-infected	-1.18	0.51	101	-2.30	0.303
	Boa Vista non-infected - Boa Vista infected	-0.32	0.45	102	-0.70	0.997
	Boa Vista non-infected - Maio infected	-1.06	0.46	102	-2.31	0.299
	Boa Vista non-infected - Sal infected	-0.54	0.44	99	-1.23	0.921
	Boa Vista non-infected - Santo Antão infected	-0.50	0.46	102	-1.10	0.956
	Maio non-infected - Sal non-infected	-0.62	0.43	102	-1.44	0.834
	Maio non-infected - Santo Antão non-infected	-1.11	0.46	101	-2.45	0.232
	Maio non-infected - Boa Vista infected	-0.25	0.39	100	-0.63	0.998
	Maio non-infected - Maio infected	-0.99	0.41	101	-2.44	0.235
	Maio non-infected - Sal infected	-0.47	0.37	101	-1.28	0.906
	Maio non-infected - Santo Antão infected	-0.43	0.40	100	-1.09	0.958
	Sal non-infected - Santo Antão non-infected	-0.50	0.41	101	-1.22	0.926
	Sal non-infected - Boa Vista infected	0.37	0.34	102	1.08	0.959
	Sal non-infected - Maio infected	-0.37	0.36	102	-1.03	0.969
	Sal non-infected - Sal infected	0.15	0.31	102	0.49	1.000
	Sal non-infected - Santo Antão infected	0.18	0.35	102	0.51	1.000
	Santo Antão non-infected - Boa Vista infected	0.86	0.38	101	2.30	0.303
	Santo Antão non-infected - Maio infected	0.12	0.40	102	0.31	1.000
	Santo Antão non-infected - Sal infected	0.65	0.34	100	1.88	0.569
	Santo Antão non-infected - Santo Antão infected	0.68	0.39	102	1.75	0.655
	Boa Vista infected - Maio infected	-0.74	0.32	102	-2.32	0.295

Boa Vista infected - Sal infected	-0.22	0.26	101	-0.84	0.991
Boa Vista infected - Santo Antão infected	-0.19	0.31	101	-0.60	0.999
Maio infected - Sal infected	0.52	0.29	100	1.80	0.619
Maio infected - Santo Antão infected	0.55	0.33	101	1.70	0.686
Sal infected - Santo Antão infected	0.03	0.28	102	0.12	1.000

Supplementary Table 5.4: Pairwise PERMANOVA results of differences in microbial community composition between turtle nesting islands for rarefied dry swab samples.

Pairwise comparisons	Df	Unweighted UniFrac			Weighted UniFrac			Bray-Curtis		
		F	R2	Pr(>F)	F	R2	Pr(>F)	F	R2	Pr(>F)
Boa Vista vs Sal	1,71	1.14	0.02	0.272	1.31	0.02	0.220	1.08	0.02	0.321
Boa Vista vs Santo Antão	**1,48**	**2.21**	**0.04**	**0.003**	1.46	0.03	0.146	**1.48**	**0.03**	**0.035**
Boa Vista vs Maio	1,51	1.24	0.02	0.145	1.40	0.03	0.169	0.94	0.02	0.560
Sal vs Santo Antão	**1,71**	**2.38**	**0.03**	**0.001**	**2.27**	**0.03**	**0.028**	**1.72**	**0.02**	**0.031**
Sal vs Maio	1,74	1.28	0.02	0.124	1.19	0.02	0.267	1.44	0.02	0.075
Santo Antão vs Maio	**1,51**	**2.06**	**0.04**	**0.014**	1.44	0.03	0.145	**1.57**	**0.03**	**0.017**

Supplementary Table 5.5:Results from SIMPER for non-rarefied dry swabs that significantly differ in abundance between turtle nesting islands. MA – Maio, SA – Santo Antão, SL – Sal and BV – Boa Vista. Table shows ASVs, their mean abundance in each island and their taxonomic information.

Comparison (X - Y)	ASV	p.value	Mean Abundance X ± SD	Mean Abundance Y ± SD	Phylum	Class	Order	Family	Genus	Species
MA - SA	81554b0fce7cb99c919725c40ec6792a	0.022	0.007 ± 0.026	0.020 ± 0.046	Bacteroidetes	Bacteroidia	Bacteroidales	[Paraprevotellaceae]	NA	NA
SL - SA	81554b0fce7cb99c919725c40ec6792a	0.032	0.006 ± 0.015	0.020 ± 0.046	Bacteroidetes	Bacteroidia	Bacteroidales	[Paraprevotellaceae]	NA	NA
BV - SA	17fbebb3d9f3bbd3508cd5348ae7ea7c	0.005	0.008 ± 0.019	0.074 ± 0.110	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	NA	NA	NA
MA - SA	17fbebb3d9f3bbd3508cd5348ae7ea7c	0.006	0.033 ± 0.081	0.074 ± 0.110	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	NA	NA	NA
SL - SA	17fbebb3d9f3bbd3508cd5348ae7ea7c	0.016	0.032 ± 0.063	0.074 ± 0.110	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	NA	NA	NA
BV - SA	0015124e5777b57174a1af7411703bbb	0.021	0.018 ± 0.037	0.004 ± 0.007	Proteobacteria	Gammaproteobacteria	Pasteurellales	Pasteurellaceae	NA	NA
BV - SA	413037c7bfae2c3535793bc67f4d9ae5	0.012	0.042 ± 0.066	0.019 ± 0.050	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	NA
MA - SA	0429a8a999c3238e12bbfaa1714d385e	0.002	0.091 ± 0.136	0.039 ± 0.067	Proteobacteria	Gammaproteobacteria	Vibrionales	NA	NA	NA
MA - SA	34239f14f57eed8a2971a240df604a06	<0.001	0.028 ± 0.065	0.000 ± 0.001	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	Shewanella	algae
BV - MA	0015124e5777b57174a1af7411703bbb	0.037	0.018 ± 0.037	0.015 ± 0.044	Proteobacteria	Gammaproteobacteria	Pasteurellales	Pasteurellaceae	NA	NA
BV - MA	0429a8a999c3238e12bbfaa1714d385e	0.019	0.133 ± 0.243	0.039 ± 0.067	Proteobacteria	Gammaproteobacteria	Vibrionales	NA	NA	NA
BV - MA	34239f14f57eed8a2971a240df604a06	0.034	0.046 ± 0.175	0.000 ± 0.001	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	Shewanella	algae
BV - MA	413037c7bfae2c3535793bc67f4d9ae5	0.017	0.042 ± 0.066	0.047 ± 0.147	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	NA
BV - SL	413037c7bfae2c3535793bc67f4d9ae5	0.006	0.042 ± 0.066	0.016 ± 0.060	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	NA
SL - MA	34239f14f57eed8a2971a240df604a06	0.005	0.019 ± 0.079	0.000 ± 0.001	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	Shewanella	algae
SL - MA	ad7f3efd185f47aab4f7badf2c822cc8	0.009	0.020 ± 0.034	0.011 ± 0.024	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	Helicobacter	NA
BV - SL	33c1fcdbc21a10a2a033f53d47db7893	0.019	0.072 ± 0.186	0.025 ± 0.074	Tenericutes	Mollicutes	Mycoplasmatales	Mycoplasmataceae	Mycoplasma	NA

Supplementary Table 5.6: Results from SIMPER for rarefied dry swabs that significantly differ in abundance between turtle nesting islands. MA – Maio, SA – Santo Antão, SL – Sal and BV – Boa Vista. Table shows ASVs, their mean abundance in each island and their taxonomic information.

Comparison (X - Y)	ASV	p.value	Mean Abundance X ± SD	Mean Abundance Y ± SD	Phylum	Class	Order	Family	Genus	Species
BV - MA	34239f14f57eed8a2971a240df604a06	0.038	0.053 ± 0.192	0.000 ± 0.000	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewanella*	*algae*
BV - SL	413037c7bfae2c3535793bc67f4d9ae5	0.003	0.045 ± 0.067	0.016 ± 0.063	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	*Campylobacter*	*NA*
BV - SA	0015124e5777b57174a1af7411703bbb	0.020	0.020 ± 0.040	0.004 ± 0.008	Proteobacteria	Gammaproteobacteria	Pasteurellales	Pasteurellaceae	*NA*	*NA*
BV - SA	17fbebb3d9f3bbd3508cd5348ae7ea7c	<0.001	0.006 ± 0.014	0.080 ± 0.111	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	*NA*	*NA*	*NA*
BV - SA	413037c7bfae2c3535793bc67f4d9ae5	0.029	0.045 ± 0.067	0.019 ± 0.049	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	*Campylobacter*	*NA*
BV - SA	d20b46e3c9d79a8e49a48f112fc03d4f	0.022	0.000 ± 0.000	0.020 ± 0.090	Firmicutes	Clostridia	Clostridiales	Peptostreptococcaceae	*NA*	*NA*
MA - SA	17fbebb3d9f3bbd3508cd5348ae7ea7c	0.007	0.034 ± 0.086	0.080 ± 0.111	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	*NA*	*NA*	*NA*
MA - SA	34239f14f57eed8a2971a240df604a06	0.001	0.000 ± 0.000	0.028 ± 0.067	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewanella*	*algae*
MA - SA	d20b46e3c9d79a8e49a48f112fc03d4f	0.025	0.000 ± 0.000	0.020 ± 0.090	Firmicutes	Clostridia	Clostridiales	Peptostreptococcaceae	*NA*	*NA*
SL - MA	34239f14f57eed8a2971a240df604a06	0.008	0.021 ± 0.082	0.000 ± 0.000	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewanella*	*algae*
SL - SA	17fbebb3d9f3bbd3508cd5348ae7ea7c	0.004	0.034 ± 0.065	0.080 ± 0.111	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	*NA*	*NA*	*NA*
SL - SA	81554b0fce7cb99c919725c40ec6792a	0.032	0.006 ± 0.015	0.021 ± 0.048	Bacteroidetes	Bacteroidia	Bacteroidales	[Paraprevotellaceae]	*NA*	*NA*
SL - SA	95064b90639c6c20f85c141b5e8d6a15	0.027	0.009 ± 0.022	0.014 ± 0.024	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	*NA*	*NA*
SL - SA	d20b46e3c9d79a8e49a48f112fc03d4f	0.031	0.000 ± 0.002	0.020 ± 0.090	Firmicutes	Clostridia	Clostridiales	Peptostreptococcaceae	*NA*	*NA*

Supplementary Table 5.7: Results from SIMPER for non-rarefied and rarefied dry swabs that significantly differ in abundance between infected with O. margoi and non-infected turtles. Table shows ASVs, their mean abundance in infected vs non-infected individuals and their taxonomic information. Infects - O. margoi present, N-infected - O. margoi not present.

Comparison (X - Y)	ASV	p.value	Mean Abundance X ± SD	Mean Abundance Y ± SD	Phylum	Class	Order	Family	Genus	Species
Non-rarefied										
Infected – N-infected	0429a8a999c3238e12bbfaa1714d385e	0.010	0.091 ± 0.163	0.040 ± 0.054	Proteobacteria	Gammaproteobacteria	Vibrionales	NA	*NA*	*NA*
Infected - N-infected	5699b09d3c91c083e8ce6bcafc742e25	0.010	0.011 ± 0.018	0.005 ± 0.014	Fusobacteria	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	*Fusobacterium*	*NA*
Infected - N-infected	76fd33f14910b1cd69086786956646b3	<0.001	0.050 ± 0.120	0.008 ± 0.037	Proteobacteria	Gammaproteobacteria	Cardiobacteriales	NA	*NA*	*NA*
Infected - N-infected	95064b90639c6c20f85c141b5e8d6a15	0.016	0.016 ± 0.042	0.007 ± 0.020	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	*NA*	*NA*
Infected - N-infected	f62dc709536edba0135034e740915a92	0.013	0.007 ± 0.025	0.044 ± 0.104	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	*Helicobacter*	*NA*
Rarefied										
Infected - N-infected	d1b122b873bef2a9246fbd7ead125ca7	0.021	0.006 ± 0.025	0.016 ± 0.048	Bacteroidetes	Flavobacteriia	Flavobacteriales	[Weeksellaceae]	*Ornithobacterium*	*NA*
Infected - N-infected	f62dc709536edba0135034e740915a92	0.013	0.007 ± 0.026	0.046 ± 0.111	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	*Helicobacter*	*NA*
Infected - N-infected	76fd33f14910b1cd69086786956646b3	0.000	0.049 ± 0.120	0.002 ± 0.005	Proteobacteria	Gammaproteobacteria	Cardiobacteriales	*NA*	*NA*	*NA*

Supplementary Table 5.8: Pairwise PERMANOVA results showing differences in microbial community composition between the nesting island of a turtle and its size for the non-rarefied wet swab dataset

Pairwise comparisons	Df	Weighted UniFrac			Bray Curtis		
		F	R2	Pr(>F)	F	R2	Pr(>F)
Sal Small vs Boa Vista Medium	1,26	0.97	0.04	0.420	1.08	0.04	0.322
Sal Small vs Santo Antão Large	1,27	0.97	0.04	0.415	1.15	0.04	0.231
Sal Small vs Sal Large	1,31	1.10	0.04	0.339	1.10	0.04	0.265
Sal Small vs Santo Antão Medium	1,21	1.25	0.06	0.230	1.11	0.05	0.324
Sal Small vs Boa Vista Large	1,23	1.53	0.06	0.141	1.24	0.05	0.143
Sal Small vs Maio Medium	1,23	1.12	0.05	0.291	1.06	0.05	0.339
Sal Small vs Santo Antão Small	1,20	0.76	0.04	0.630	1.49	0.07	0.067
Sal Small vs Maio Small	1,22	1.12	0.05	0.295	1.19	0.05	0.188
Sal Small vs Sal Medium	1,26	1.34	0.05	0.224	1.27	0.05	0.169
Sal Small vs Maio Large	1,24	1.20	0.05	0.239	1.00	0.04	0.450
Sal Small vs Boa Vista Small	1,19	1.31	0.07	0.223	1.12	0.06	0.281
Boa Vista Medium vs Santo Antão Large	1,20	1.76	0.08	0.099	1.21	0.06	0.159
Boa Vista Medium vs Sal Large	1,24	0.48	0.02	0.917	0.93	0.04	0.513
Boa Vista Medium vs Santo Antão Medium	1,14	0.65	0.05	0.777	0.76	0.06	0.775
Boa Vista Medium vs Boa Vista Large	1,16	0.62	0.04	0.789	0.92	0.06	0.581
Boa Vista Medium vs Maio Medium	1,16	0.51	0.03	0.868	0.75	0.05	0.778
Boa Vista Medium vs Santo Antão Small	1,13	1.40	0.10	0.156	**1.81**	**0.13**	**0.012**
Boa Vista Medium vs Maio Small	1,15	0.55	0.04	0.851	0.72	0.05	0.838
Boa Vista Medium vs Sal Medium	1,19	1.26	0.07	0.225	1.34	0.07	0.095
Boa Vista Medium vs Maio Large	**1,17**	**2.35**	**0.13**	**0.032**	1.34	0.08	0.123
Boa Vista Medium vs Boa Vista Small	1,12	0.70	0.06	0.776	0.78	0.07	0.797
Santo Antão Large vs Sal Large	1,25	1.89	0.07	0.069	1.10	0.04	0.304
Santo Antão Large vs Santo Antão Medium	**1,15**	**2.72**	**0.16**	**0.024**	1.40	0.09	0.082
Santo Antão Large vs Boa Vista Large	**1,17**	**2.43**	**0.13**	**0.031**	1.35	0.08	0.081
Santo Antão Large vs Maio Medium	1,17	1.61	0.09	0.120	1.18	0.07	0.192
Santo Antão Large vs Santo Antão Small	1,14	0.58	0.04	0.824	1.17	0.08	0.250
Santo Antão Large vs Maio Small	1,16	1.85	0.11	0.065	1.16	0.07	0.231
Santo Antão Large vs Sal Medium	**1,20**	**2.65**	**0.12**	**0.025**	**1.62**	**0.08**	**0.040**
Santo Antão Large vs Maio Large	1,18	1.33	0.07	0.245	1.00	0.06	0.444
Santo Antão Large vs Boa Vista Small	1,13	2.00	0.14	0.068	1.18	0.09	0.171
Sal Large vs Santo Antão Medium	1,19	1.12	0.06	0.332	0.90	0.05	0.632
Sal Large vs Boa Vista Large	1,21	1.01	0.05	0.414	1.38	0.06	0.056
Sal Large vs Maio Medium	1,21	0.36	0.02	0.980	0.93	0.04	0.540
Sal Large vs Santo Antão Small	1,18	1.45	0.08	0.153	**1.55**	**0.08**	**0.048**
Sal Large vs Maio Small	1,20	0.47	0.02	0.944	0.99	0.05	0.446
Sal Large vs Sal Medium	**1,24**	**1.86**	**0.07**	**0.042**	1.41	0.06	0.091
Sal Large vs Maio Large	**1,22**	**3.07**	**0.13**	**0.013**	1.45	0.06	0.064
Sal Large vs Boa Vista Small	1,17	1.19	0.07	0.267	1.18	0.07	0.193
Santo Antão Medium vs Boa Vista Large	1,11	0.72	0.07	0.727	1.26	0.11	0.134
Santo Antão Medium vs Maio Medium	1,11	1.03	0.09	0.432	0.98	0.09	0.486
Santo Antão Medium vs Santo Antão Small	**1,8**	**2.36**	**0.25**	**0.041**	1.97	0.22	0.053
Santo Antão Medium vs Maio Small	1,10	1.31	0.13	0.275	0.98	0.10	0.472
Santo Antão Medium vs Sal Medium	1,14	0.69	0.05	0.770	0.99	0.07	0.423
Santo Antão Medium vs Maio Large	1,12	2.22	0.17	0.089	1.42	0.11	0.120
Santo Antão Medium vs Boa Vista Small	1,7	0.80	0.12	0.634	1.20	0.17	0.225
Boa Vista Large vs Maio Medium	1,13	0.61	0.05	0.752	1.01	0.08	0.444
Boa Vista Large vs Santo Antão Small	1,10	2.12	0.19	0.073	**1.94**	**0.18**	**0.004**
Boa Vista Large vs Maio Small	1,12	1.31	0.11	0.223	1.14	0.09	0.334
Boa Vista Large vs Sal Medium	1,16	1.04	0.06	0.388	1.14	0.07	0.265
Boa Vista Large vs Maio Large	**1,14**	**2.32**	**0.15**	**0.048**	1.28	0.09	0.140
Boa Vista Large vs Boa Vista Small	1,9	0.97	0.11	0.387	0.86	0.10	0.704

Maio Medium vs Santo Antão Small	1,10	1.38	0.13	0.216	1.40	0.13	0.101
Maio Medium vs Maio Small	1,12	0.68	0.06	0.714	0.85	0.07	0.706
Maio Medium vs Sal Medium	1,16	1.62	0.10	0.088	1.17	0.07	0.243
Maio Medium vs Maio Large	**1,14**	**2.31**	**0.15**	**0.049**	1.24	0.09	0.161
Maio Medium vs Boa Vista Small	1,9	1.04	0.11	0.427	0.85	0.10	0.648
Santo Antão Small vs Maio Small	1,9	2.09	0.21	0.074	**1.74**	**0.18**	**0.028**
Santo Antão Small vs Sal Medium	1,13	1.90	0.14	0.052	**1.80**	**0.13**	**0.045**
Santo Antão Small vs Maio Large	1,11	0.88	0.08	0.433	0.95	0.09	0.471
Santo Antão Small vs Boa Vista Small	1,6	1.84	0.27	0.127	1.91	0.28	0.052
Maio Small vs Sal Medium	**1,15**	**1.83**	**0.12**	**0.028**	1.23	0.08	0.195
Maio Small vs Maio Large	**1,13**	**2.97**	**0.20**	**0.018**	**1.53**	**0.11**	**0.019**
Maio Small vs Boa Vista Small	1,8	1.16	0.14	0.302	1.17	0.14	0.247
Sal Medium vs Maio Large	1,17	1.57	0.09	0.135	1.12	0.07	0.306
Sal Medium vs Boa Vista Small	1,12	1.08	0.09	0.303	1.07	0.09	0.358
Maio Large vs Boa Vista Small	1,10	1.76	0.16	0.159	1.10	0.11	0.391

Supplementary Table 5.9: Results from SIMPER for non-rarefied wet swabs that significantly differ in abundance with an interaction between turtle nesting island and its size. Table shows ASVs, their mean abundance amongst pairs and their taxonomic information. L – large sized turtles, M- Medium sized turtles and S- Small sized turtles.

Comparison (X - Y)	ASV	p.value	Mean Abundance X ± SD	Mean Abundance Y ± SD	Phylum	Class	Order	Family	Genus	Species
Boa Vista										
L-S	5699b09d3c91c083e8ce6bcafc742e25	0.050	0.004 ± 0.009	0.019 ± 0.026	Fusobacteria	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	*Fusobacterium*	-
M-L	08b7ae36ca8ff80a868794ff4aa3af1c	0.027	0.002 ± 0.005	0.013 ± 0.027	Actinobacteria	Actinobacteria	-	-	-	-
M-S	9be79137deae3d1b144b96c452427950	0.037	0.006 ± 0.007	0.015 ± 0.011	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*	-
Sal										
L-M	0429a8a999c3238e12bbfaa1714d385e	0.012	0.059 ± 0.049	0.009 ± 0.016	Proteobacteria	Gammaproteobacteria	Vibrionales	-	-	-
S-M	0429a8a999c3238e12bbfaa1714d385e	0.007	0.063 ± 0.095	0.009 ± 0.016	Proteobacteria	Gammaproteobacteria	Vibrionales	-	-	-
S-M	21e65a24850031f756e25f0278d4d1af	0.028	0.023 ± 0.042	0.002 ± 0.005	Bacteroidetes	Bacteroidia	Bacteroidales	-	-	-
Maio										
M-L	95064b90639c6c20f85c141b5e8d6a15	0.016	0.000 ± 0.000	0.038 ± 0.059	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	-	-
M-L	ff0b033d6bfeaf9f7afadb421c67cfbc	0.038	0.000 ± 0.000	0.024 ± 0.041	Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae	*Parabacteroides*	-
S-L	9be79137deae3d1b144b96c452427950	0.027	0.003 ± 0.008	0.019 ± 0.020	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*	-
S-L	f0b70acd270f53b521eef73b3128a67f	0.024	0.000 ± 0.000	0.056 ± 0.117	Bacteroidetes	Bacteroidia	Bacteroidales	Rikenellaceae	*AF12*	-
Santo Antão										
L-M	17fbebb3d9f3bbd3508cd5348ae7ea7c	0.033	0.075 ± 0.084	0.003 ± 0.006	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	-	-	-
L-M	95064b90639c6c20f85c141b5e8d6a15	0.038	0.020 ± 0.035	0.002 ± 0.004	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	-	-
L-M	9be79137deae3d1b144b96c452427950	0.012	0.020 ± 0.016	0.002 ± 0.005	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*	-
L-M	ce7504bf8df2af8e3139aacad6b79dae	0.044	0.021 ± 0.021	0.005 ± 0.010	Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae	*Paludibacter*	-
L-M	f0b70acd270f53b521eef73b3128a67f	0.029	0.015 ± 0.031	0.000 ± 0.000	Bacteroidetes	Bacteroidia	Bacteroidales	Rikenellaceae	*AF12*	-
M-S	17fbebb3d9f3bbd3508cd5348ae7ea7c	0.013	0.003 ± 0.006	0.164 ± 0.113	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	-	-	-
M-S	1aeef1011c39dc1e7b526ebeab794a68	0.021	0.002 ± 0.004	0.042 ± 0.020	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae	-	-
M-S	81554b0fce7cb99c919725c40ec6792a	0.046	0.001 ± 0.002	0.018 ± 0.020	Bacteroidetes	Bacteroidia	Bacteroidales	[Paraprevotellaceae]	-	-
M-S	9be79137deae3d1b144b96c452427950	0.011	0.002 ± 0.005	0.032 ± 0.017	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*	-
M-S	ce7504bf8df2af8e3139aacad6b79dae	0.041	0.005 ± 0.010	0.028 ± 0.029	Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae	*Paludibacter*	-

Supplementary Table 5.10: Results from SIMPER for non-rarefied wet swabs that significantly differ in abundance between infected with O. margoi and non-infected turtles. Table shows ASVs, their mean abundance in infected vs non-infected individuals and their taxonomic information. Infects - O. margoi present, N-infected - O. margoi not present.

Comparison (X - Y)	ASV	p.value	Mean Abundance X ± SD	Mean Abundance Y ± SD	Phylum	Class	Order	Family	Genus	Species
Infected_N-infected	17fbebb3d9f3bbd3508cd5348ae7ea7c	0.010	0.058 ± 0.098	0.020 ± 0.057	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	-	-	-
Infected_N-infected	62312a4c1b699acbc6a6862c1a8bea83	0.006	0.030 ± 0.071	0.019 ± 0.058	Fusobacteria	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	NA	NA
Infected_N-infected	76fd33f14910b1cd69086786956646b3	0.000	0.044 ± 0.107	0.008 ± 0.039	Proteobacteria	Gammaproteobacteria	Cardiobacteriales	NA	NA	NA
Infected_N-infected	ce7504bf8df2af8e3139aacad6b79dae	0.016	0.020 ± 0.038	0.018 ± 0.042	Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae	Paludibacter	NA
Infected_N-infected	21e65a24850031f756e25f0278d4d1af	0.035	0.012 ± 0.025	0.009 ± 0.024	Bacteroidetes	Bacteroidia	Bacteroidales	NA	NA	NA

Supplementary Table 5.11: Results from SIMPER for rarefied wet swabs that significantly differ in abundance between turtles nesting on different islands. Table shows ASVs, their mean abundance in infected vs non-infected individuals and their taxonomic information. MA – Maio, SA – Santo Antão, SL – Sal and BV – Boa Vista.

Comparison (X - Y)	ASV	p.value	Mean Abundance X ± SD	Mean Abundance Y ± SD	Phylum	Class	Order	Family	Genus	Species
BV_MA	76fd33f14910b1cd69086786956646b3	0.03519	0.046 ± 0.112	0.006 ± 0.020	Proteobacteria	Gammaproteobacteria	Cardiobacteriales	NA	NA	NA
BV_SA	0015124e5777b57174a1af7411703bbb	0.009613	0.021 ± 0.039	0.002 ± 0.006	Proteobacteria	Gammaproteobacteria	Pasteurellales	Pasteurellaceae	NA	NA
BV_SA	17fbebb3d9f3bbd3508cd5348ae7ea7c	0.021726	0.004 ± 0.005	0.104 ± 0.111	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	NA	NA	NA
BV_SA	67b0cf6af26b9fd959d860a52a65bc58	0.035141	0.002 ± 0.003	0.071 ± 0.117	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	Photobacterium	damselae
BV_SA	81554b0fce7cb99c919725c40ec6792a	0.001356	0.001 ± 0.002	0.025 ± 0.031	Bacteroidetes	Bacteroidia	Bacteroidales	[Paraprevotellaceae]	NA	NA
SL_BV	413037c7bfae2c3535793bc67f4d9ae5	0.021065	0.010 ± 0.028	0.025 ± 0.027	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	NA
SL_BV	67b0cf6af26b9fd959d860a52a65bc58	0.010218	0.029 ± 0.040	0.002 ± 0.003	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	Photobacterium	damselae
SL_MA	67b0cf6af26b9fd959d860a52a65bc58	0.021516	0.029 ± 0.040	0.003 ± 0.007	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	Photobacterium	damselae
SL_SA	0429a8a999c3238e12bbfaa1714d385e	0.040526	0.053 ± 0.086	0.065 ± 0.054	Proteobacteria	Gammaproteobacteria	Vibrionales	NA	NA	NA
SL_SA	1aeef1011c39dc1e7b526ebeab794a68	0.025382	0.007 ± 0.019	0.018 ± 0.022	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae	NA	NA

SL_SA	9be79137deae3d1b144b96c452427950	0.031645	0.009 ± 0.014	0.019 ± 0.016	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	Vibrio	*NA*
SL_SA	b501af89d0069bb6642c88a656b4d962	0.02781	0.022 ± 0.045	0.001 ± 0.004	Proteobacteria	Betaproteobacteria	Neisseriales	Neisseriaceae	*NA*	*NA*

Supplementary Table 5.12: Results from SIMPER for rarefied wet swabs that significantly differ in abundance between infected with O. margoi and non-infected turtles. Table shows ASVs, their mean abundance in infected vs non-infected individuals and their taxonomic information. Infects - O. margoi present, N-infected - O. margoi not present.

Comparison (X - Y)	*ASV*	*p.value*	*Mean Abundance X ± SD*	*Mean Abundance Y ± SD*	Phylum	Class	Order	Family	*Genus*	*Species*
Infected - N-infected	17fbebb3d9f3bbd3508cd5348ae7ea7c	0.006	0.077 ± 0.116	0.010 ± 0.032	Lentisphaerae	[Lentisphaeria]	Lentisphaerales	*NA*	*NA*	*NA*
Infected - N-infected	5699b09d3c91c083e8ce6bcafc742e25	0.028	0.016 ± 0.022	0.011 ± 0.028	Fusobacteria	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium	*NA*
Infected - N-infected	62312a4c1b699acbc6a6862c1a8bea83	0.023	0.041 ± 0.088	0.030 ± 0.079	Fusobacteria	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	*NA*	*NA*
Infected - N-infected	76fd33f14910b1cd69086786956646b3	0.002	0.031 ± 0.077	0.000 ± 0.001	Proteobacteria	Gammaproteobacteria	Cardiobacteriales	*NA*	*NA*	*NA*
Infected - N-infected	b4bb370a59f2637d406ccc0f5a1b2a1e	0.049	0.021 ± 0.111	0.000 ± 0.000	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Morganella	morganii
Infected - N-infected	ce7504bf8df2af8e3139aacad6b79dae	0.039	0.024 ± 0.048	0.010 ± 0.024	Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae	Paludibacter	*NA*