

Environmental and Genetic Determinants
of Bacterial Microbiota and Parasites
Associated with *Cottus* Across Natural Hybrid Zones

Dissertation

in fulfillment of the requirement for the degree

“Dr. rer. nat.”

of the Faculty of Mathematics and Natural Sciences

at the Christian Albrechts University of Kiel

submitted by

Sunna Ellendt

Kiel, July 2016

First referee:

Prof. Dr. Arne W. Nolte

Second referee:

Prof. Dr. Thorsten Reusch

Date of the oral examination:

12.10.2016

Approved for publication:

12.10.2016

Prof. Dr. Natascha Oppelt, Dean

TABLE OF CONTENTS

Table of Contents i

Zusammenfassung iii

Summary v

Author contributions vii

General Introduction 1

 Speciation 1

 Parasite induced reproductive isolation 3

 Bacteria induced reproductive isolation 5

 Hybrid zones as natural laboratories to study early adaptive speciation 8

 The *Cottus* System 9

Scope of the thesis 12

Chapter 1 13

 Introduction 13

 Methods 16

 Results 22

 Discussion 35

 Supplement 41

Chapter 2 51

 Introduction 51

 Methods 53

 Results 56

 Discussion 65

 Supplement 70

Chapter 3 77

 Introduction 77

 Methods 80

 Results 83

 Supplement 90

Conclusions and perspectives 93

 Particular focus on host genetic effects 93

 Environmental impact on bacterial communities 94

 Relation of host-pathogen interactions to adaptive immune response 95

 Conclusion 95

Table of Contents

References	97
Danksagung	123
Affidavit	125

ZUSAMMENFASSUNG

Wie Arten entstehen ist ein bedeutender evolutionärer Prozess und für viele Wissenschaftler von großem Interesse (u.a. Schluter 2001; Coyne & Orr 2004; Nosil *et al.* 2005; Karvonen & Seehausen 2012). Neben genetischen (intrinsischen) Isolationsmechanismen können sich auch nicht-genetische (extrinsische), ökologisch begründete Fortpflanzungsbarrieren durch divergierende Selektion zwischen verschiedenen Umwelten entwickeln (Rundle & Nosil 2005). Zugrunde liegende biotische und abiotische Faktoren beinhalten auch Interaktionen mit Symbionten, die bedeutenden Einfluss auf ihren Wirt ausüben: Sie können ihm einerseits Nährstoffe oder auch Abwehr gegen Pathogene bereitstellen und ihn andererseits ausbeuten oder infizieren (Moran 2006). Jedoch wurde Bakterien und Parasiten als selektierende Faktoren in Artbildungsprozessen bisher wenig Aufmerksamkeit zuteil und insbesondere Studien über freilebende Wirtspopulationen fehlen (Karvonen & Seehausen 2012; Brucker & Bordenstein 2012). In der vorliegenden Arbeit untersuchte ich eine natürlich entstandene Kontaktzone europäischer Süßwasserfische (*Cottus* sp.), die an einem Ökoton hybridisieren. Diese Hybridzone ist stabil und obwohl keine physikalischen Barrieren vorhanden sind, ist die Ausbreitung der Fische in fremde Habitats begrenzt (Nolte *et al.* 2005). Darüber hinaus wurden keine intrinsischen Barrieren entdeckt weshalb davon ausgegangen werden kann, dass extrinsische, ökologische Faktoren Selektionsdruck auf Migranten ausüben und deren verminderte Fitness in alternativen Nischen verursachen (Nolte *et al.* 2006; Cheng *et al.* 2013). Die vorliegende Studie erforscht also, in welchem Ausmaß Bakterien und Parasiten anfängliche Artbildungsprozesse im *Cottus* System vorantreiben.

Basierend auf früheren populationsgenetischen Untersuchungen definierte ich zunächst drei unterscheidbare Fischgruppen, welche die aus dem Fluss Sieg und dem Bach Broel gebildete Kontaktzone bewohnen: *C. rhenanus*, invasive *C.* und deren Hybride. Anschließend analysierte ich die Hautbakteriengemeinschaften von *Cottus*. Das Ziel war herauszufinden, ob Variationen in den Bakterien durch Unterschiede in der Fischgenetik und / oder im -habitat beeinflusst werden. Mithilfe von Hochdurchsatzsequenzierung des 16S rRNA Markergens zeigte sich ein starker Habitatsinfluss auf die bakterielle Diversität in den unterschiedlichen Fischlinien und ein geringerer wirtsgenetischer Effekt. Um für umweltbedingte Varianz zu kontrollieren, untersuchte ich danach die bakterielle Zusammensetzung von *C. rhenanus* und invasive *C.*, die im Labor unter gleichartigen Haltungsbedingungen aufgezogen wurden. Die Ergebnisse einiger Analysen deuteten auf einen wirtsgenetischen Einfluss auf die bakterielle Zusammensetzung hin. Desweiteren verglich ich die bakterielle Mikrobiota von freilebenden und im Labor aufgezogene Gruppen miteinander und ermittelte, trotz beträchtlicher Unterschiede in Futter und Habitat, eine ähnliche Auftrennung nach Fischgenotyp in beiden Gruppen, welche zum Teil unabhängig von der umgebenden Umwelt war.

Um wirtsgenetische Faktoren von Umweltfaktoren vollständig zu entkoppeln, analysierte ich anschließend die Hautbakterien anderer Knochenfische, *Barbatula barbatula* und

Zusammenfassung

Phoxinus phoxinus. Diese Arten kommen in beiden Habitaten der Kontaktzone in großer Anzahl vor, wobei Fluss- und Bachtiere genetisch homogen sind. Die Ergebnisse zeigten einen bedeutenden wirtsgenetischen Effekt an: Die Varianz der Hautbakterien in unterschiedlichen Fischarten aus demselben Habitat war größer, als die derselben Art in verschiedenen Habitaten. Darüber hinaus verglich ich die Bakteriengemeinschaften von *C. rhenanus* und invasive *C.*, gefangen in 2005 und 2014, und ermittelte einen ähnlichen wirtsgenetischen Effekt in beiden Sammlungsjahren. Dieses spiegelt vorherige Ergebnisse wider und deutet auf eine Kern-Mikrobiota in *Cottus* hin.

SUMMARY

How species evolve is a fundamental evolutionary process which has been of great interest to many scientists (e.g. Schluter 2001; Coyne and Orr 2004; Nosil *et al.* 2005; Nosil 2012; Seehausen *et al.* 2014). Besides genetic (intrinsic) isolating mechanisms, barriers to gene flow between populations can emerge as a result of non-genetic (extrinsic), ecologically-based divergent selection between different environments (Rundle and Nosil 2005). Underlying biotic and abiotic factors also include interactions with symbionts having a significant influence on their hosts: they provide nutrients or defense against pathogens on the one hand but can exploit or infect them on the other hand (Moran 2006). However, bacteria and parasites have gained little attention as selective agents in speciation processes (Brucker and Bordenstein 2012; Karvonen and Seehausen 2012) and particularly studies in wild host populations are lacking. In this work I analyzed a naturally emerged contact zone of European freshwater sculpins (*Cottus* sp.) hybridizing at a habitat ecotone. The hybrid zone is stable and despite the absence of physical barriers dispersal of the fish to foreign habitats is limited (Nolte *et al.* 2005). Further, no intrinsic barriers have been revealed, thus it has been suggested that extrinsic ecological factors exert selective pressure on migrants and reduce their fitness in alternative niches (Nolte *et al.* 2006; Cheng *et al.* 2013). Hence, this study aimed at investigating to what extent bacteria and parasites act as a driver to early adaptive speciation processes in the *Cottus* system.

Based on previous population genetic analysis I first defined three distinguishable fish groups: *C. rhenanus*, invasive *C.* and their hybrids inhabiting the contact zone river Sieg / stream Broel. Subsequently, I examined the skin bacterial community structure in *Cottus*. The goal was to characterize whether bacterial variation is influenced by fish genetic and / or habitat differences. By applying high throughput sequence analysis of the bacterial 16S rRNA marker gene I identified a strong influence of habitat on bacterial diversity in the different fish lineages and a minor effect of host genotype. To control for environmental variation I then investigated the bacterial composition of *C. rhenanus* and invasive *C.* being kept in the laboratory under similar rearing conditions. The results suggested an influence of host genotype on the bacterial composition in some measurements. Further, I compared the bacterial microbiota of wild and lab-raised *Cottus* and despite their considerable difference concerning food and habitat conditions I identified a similar separation between the genotypes in both groups, which is in part independent from the surrounding environment.

To completely disentangle environmental from genetic factors I subsequently investigated the skin bacteria of other teleost fish, *Barbatula barbatula* and *Phoxinus phoxinus*. These species are very abundant in both contact zone habitats and stream and river fish are genetically homogeneous. Results indicated a main effect of host genotype: the variation of skin bacteria in distinct fish species from the same habitat was larger than that of the same fish species in different habitats. Additionally, I compared the bacterial communities between wild *C. rhenanus* and invasive *C.* from 2005 and 2014 and revealed a similar host

genetic divergence within both sampling years. This mirrors previous results and suggests a core microbiota in *Cottus*.

Besides bacterial microbiota, I also investigated whether parasite abundance and diversity varies in distinct host species and families of *Cottus* by carrying out three pilot studies. First, I characterized the parasites of wild caught *C. rhenanus* and invasive *C.* and found significant effects of host genotype and body size (on average larger fish contained more parasites) on parasite abundance. Next, I exposed several lab bred families of *C. perifretum* (river Scheldt system, Netherlands), *C. rhenanus* and invasive *C.* to the common fish parasite *Diplostomum pseudospathaceum* and counted their parasite infections after 24 h. This resulted in significant differences in parasite abundance between the fish families which were not related to host genotype.

In sum, considering all results I cannot conclusively confirm that bacteria and parasites play a main role in the emergence of reproductive barriers between distinct sculpin genotypes. However, I clearly demonstrated that symbiont abundance and diversity is not only depending on host genotype but also on host habitat effects. Thus, I suggest that symbiosis is, amongst other factors still to be explored, definitely involved in adaptive speciation processes in sculpins.

AUTHOR CONTRIBUTIONS

This thesis comprises my dissertation in form of unpublished manuscripts. The design of the study was developed together with my supervisor Prof. Dr. Arne W. Nolte. The major parts of the practical laboratory work and the data analyses were mainly done by me. Heike Harre and Ralf Schmuck supported breeding and maintaining the fish stocks. Specific contributions were as follows:

Chapter I

The wild sculpins from the contact zone were caught by Arne W. Nolte in 2005. The samples were used within the framework of an unpublished Master thesis (Jendrik Schröder 2009). DNA extractions, SNP typing and population genetic analysis were done by Jendrik Schröder. Library preparations for sequencing were carried out by me supported by Nicole Thomsen. The samples were sequenced by Dr. Sven Künzel (MPI Plön) and subsequent sequence processing and data analysis were done by myself. Fish for the common garden experiment were bred by me together with Arne Nolte and Ralf Schmuck. The experiment was carried out by me. The animals were sacrificed and phenotypically classified by me; I further did the DNA extractions and library preparation. The sequencing was done by Sven Künzel. I performed sequence processing and data analysis and wrote the manuscript with contributions from Arne Nolte.

Chapter II

Fish were caught and sacrificed by Arne W. Nolte, Dr. Fabian Herder (Zoologisches Forschungsmuseum Alexander Koenig, Bonn), Dr. Jie Cheng (MPI Plön) and myself. Sampling, DNA extraction and library preparations were carried out by me and the samples were sequenced by Sven Künzel. Sequence processing and data analysis were done by me and I wrote the manuscript with contributions from Arne W. Nolte.

Chapter III

Wild sculpins were caught by Arne W. Nolte and sacrificed by me. I dissected the fish and performed the parasite screen and classification with support from Dr. Martin Kalbe (MPI Plön). Laboratory fish were maintained by Arne W. Nolte, Ralf Schmuck and me, and the infection experiments were carried out by Martin Kalbe. The data analysis was done by me and I wrote the manuscript with contributions from Arne W. Nolte.

Als Betreuer bestätige ich die Richtigkeit der oben stehenden Angaben.



GENERAL INTRODUCTION

At least since Darwin's 'On the Origin of Species by Means of Natural Selection' (1859) and shortly thereafter Wallace' 'The Malay Archipelago' (1869) evolutionary biologists aim to investigate the onset of speciation and the adaptation of species to different environments. In particular, many scientists focused on the emergence of isolating barriers affected by divergent environments, defined as ecological speciation (Schluter 2001; Rundle & Nosil 2005; Nosil 2012). However, although an enormous amount of research has been conducted on this topic, several aspects still remain unknown. Despite their indisputable significance and the impact they exert on their host's fitness, microbes and parasites have gained little attention in speciation research for a long time (Karvonen & Seehausen 2012; Brucker & Bordenstein 2012). Hence, after first introducing speciation and presenting the major factors leading to reproductive isolation, I will focus on symbiont-induced speciation, or precisely on speciation caused by parasites and bacteria. Subsequently, I will address the usage of hybrid zones as natural laboratories in speciation research and will finally introduce European freshwater sculpins (*Cottus* spp.) as a model system.

Speciation

During the last 150 years many species concepts arose, according to Coyne & Orr (2004) there are at least 25, and several remain controversial. However, the modern evolutionary synthesis, developed in the 1940s with Dobzhansky (1937) and Mayr (1942), is still the prevailing view. It merges the work of population geneticists, naturalists, zoologists, botanists and paleontologists and bridges the gap between Darwin's idea and new scientific discoveries like DNA. Thereafter, the evolution of reproductive isolating mechanisms to gene flow were considered as being the important evolutionary forces to speciation (Schluter 2001; Coyne & Orr 2004; Rundle & Nosil 2005; Lowry *et al.* 2008a; Nosil & Schluter 2011; Seehausen *et al.* 2014). Isolating barriers can emerge through divergent ecological or sexual selection (extrinsic reproductive isolation) or through genetic drift or genetic incompatibilities (intrinsic reproductive isolation) (Schluter 2000; Coyne & Orr 2004; Nosil *et al.* 2005; Seehausen *et al.* 2014). Further, reproductive isolation can arise before fertilization (prezygotic) or after (postzygotic). Although speciation research was intensified during the last 20 years some basic questions remain largely unanswered, i.e. which factors influence the evolution of reproductive barriers between diverging populations and what are the underlying genes. Ecological speciation and mutation order speciation have been defined to represent the two categories of selection (Schluter 2009). Whereas the first category is constituted by diverging environmental selective pressure, the latter arises when the environment's selective pressures are similar or identical and reproductive isolation evolves by fixation of different alleles (Mani & Clarke 1990).

Extrinsic barriers to reproduction

Extrinsic reproductive isolation causes reduced fitness or viability in migrants or hybrids (here offspring resulting from mating of two genetically distinct individuals), as response to environmental variation. That means there's gene-by-environment interactions (G x E). Possible divergent factors include habitat, season, environmental induced sexual selection and the interaction with potential mating partners (Schluter 2000, 2001; Rundle & Nosil 2005; Nosil *et al.* 2005; Seehausen *et al.* 2014).

Usually, different habitat preferences cause prezygotic isolation when dispersal is low. Bolnick *et al.* (2009) described an example for habitat isolation in two distinct ecotypes of the threespine stickleback (*Gasterosteus aculeatus*). With a classical transplant-recapture study the authors revealed a strong native habitat preference with an adaptive divergence being two- to fivefold higher relative to expected random movement. Further, they found that dispersal of the sticklebacks was phenotype-dependent: stream fish moving into the lake were morphologically more lake-like and vice versa.

Seasonal isolation arises when populations reach maturation at different seasons or the breeding periods are at different periods in the year. Two distinct forms of the yellow monkey flower (*Mimulus guttatus*), the coastal perennial and the inland annual form, flower time-delayed and are nearly completely reproductively isolated (Lowry *et al.* 2008b). Additionally, the late flowering coastal form is not able to successfully immigrate to the inland because of summer drought. By contrast, low salt tolerance prevents the inland form to immigrate to the coast so there's a general selection against immigrants.

Despite its strength, immigrant inviability is a rather newly described prezygotic barrier (Nosil *et al.* 2005; Lowry *et al.* 2008b; Nosil 2012). It means that migrants between populations exhibit decreased survival in foreign habitats because they are less adapted and as a consequence there's reduced mating. Ecologically-dependent postzygotic isolation might be confused with immigrant inviability but here divergent selection is affecting the hybrids and not the migrating parents. Hybrids are intermediate, thus there's no habitat they are adapted to. It is rather difficult to disentangle intrinsic genetic incompatibilities and ecologically-dependent hybrid fitness; however, reciprocal transplant experiments enable to detect environment-specific genetic effects. Therein, F1 hybrids between two distinct parental populations are generated. Subsequently, both parents and hybrids are reciprocally tested in either of the parental environment. Via *et al.* (2000) demonstrated that hybrids of two races of pea aphids (*Acyrtosiphon pisum*) showed significantly increased fitness than the respective parent in each of the two environments. By using artificial crossing in the laboratory it has been demonstrated that hybrids between sympatric benthic and limnetic ecotypes of threespine sticklebacks are fertile (Hatfield & Schluter 1999). In contrast, a study using F1 backcrosses in enclosures exposed to both parental habitats identified strong environmental-dependent fitness: benthic backcrossed sticklebacks outperformed limnetic backcrosses in the littoral zone whereas limnetic backcrosses grew twice as fast as benthic ones in the open water (Rundle 2002).

Ecological speciation also includes sexual selection, for instance when mating success is dependent on a trait which is at the same time linked to environmental conditions (Coyne & Orr 2004; Maan & Seehausen 2011). In contrast, sexual selection independent from the environment includes for instance exaggerated male ornamentation resulting from female preference, known as Fisher's runaway (Lande 1981). Causes for environment-dependent sexual selection are food resources, parasites, predators, as well as abiotic factors like climate. An example for predator induced isolation can be found in the Trinidadian guppy (*Poecilia reticulata*). The coloration of males is more conspicuous in upstream river areas with few predators compared to river sections with high predation rate (Schwartz & Hendry 2007). In vertebrates the major histocompatibility (MHC) complex can be affected by divergent sexual selection. MHC molecules bind pathogen-derived peptides and present them to T-cells, a type of white blood cells playing an important role in the adaptive immune system. These T-cells then recognize the pathogen and can fight against it (Janeway 2005). Several studies revealed that mate choice may induce an increase in resistance alleles due to a fitness advantage in the offspring. Indeed, it has been shown e.g. for threespine sticklebacks (Eizaguirre *et al.* 2009) that, referring to their own MHC profile, females prefer to mate with males sharing an intermediate MHC diversity. Moreover, the presence of a specific MHC haplotype was directly related to larger body size and higher resistance in males.

Intrinsic barriers to reproduction

Intrinsic reproductive isolation causes reduced fitness or viability in migrants or hybrids independent of the environment and is mediated by gene-by-gene interactions (Nosil & Schluter 2011; Seehausen *et al.* 2014). It can be induced by chromosome rearrangements, karyotype changes, gene transposition and reciprocal loss as well as sequence divergence and variation in transposons or Bateson-Dobzhansky-Muller incompatibilities (BDMI) (Maheshwari & Barbash 2011). BDMI describes that hybrid incompatibility is caused by change in more than only one gene locus (Johnson 2008). Consider an ancestral population with the genotype AABB. This population is split and in one population A evolves into *a* and in the other one B evolves into *b* (Wu & Ting 2004). Both alleles are incompatible when placed on the alternative background but don't suffer from reduced fitness in their own genetic background.

Parasite induced reproductive isolation

After the common definition by Price (1986), a parasite is an organism that lives in or on another living organism, obtains from it part or all of its organic nutrients and causes some degree of damage to its host. Parasitic species outnumber non-parasitic species, parasitism is the most prevalent 'lifestyle' on earth (Windsor 1998) and presumably every individual encounters a parasite in its life (Poulin 1996). Parasites exploit their hosts in various ways and alter their behavior, physiology and fitness (Price *et al.* 1986; Wood *et al.* 2007). Natu-

ral selection acts on parasite infectivity and also on host resistance (Carius *et al.* 2001). An evolutionary advantage of one antagonist often comes along with a disadvantage for the other antagonist, thus, co-evolutionary processes may be induced (Ebert 2008). The parasite and the host have to constantly adapt and evolve to reach their aim: maximizing their reproductive fitness. This pattern is known as 'The Red Queen hypothesis' (van Valen 1973). One form of antagonistic co-evolution is negative frequency dependent selection, i.e. the rare phenotype is favored and selection acts against the common host or parasite genotype in a time lagged frequency (Ebert 2008).

Biotic and abiotic factors in an environment can change dramatically within few meters, as can selective pressures. Hence, organisms are often adapted to their local environment meaning that a parasite's mean fitness is higher in its original than in foreign habitats and accordingly, hosts are less susceptible to parasites they are constantly exposed to (e.g. Kaltz & Shykoff 1998; Lively & Dybdahl 2000; Kawecki & Ebert 2004). That has been, for example, demonstrated in the threespine stickleback. Sticklebacks from lakes, wherein the abundance of the parasite *Diplostomum pseudospathaceum* was high, had a lower infection rate compared to those from rivers where the parasite was less abundant (Kalbe & Kurtz 2006). Further, lake and river sticklebacks being repeatedly experimentally exposed to *D. pseudospathaceum* were less infected than naïve fish (Scharsack & Kalbe 2014). The authors suggest that they apparently rapidly acquired an immune response by an increased activation of cells of the innate immune system.

Generally, several studies examining the role of parasites in speciation processes have been published and many of them focus on fish (reviewed by Karvonen & Seehausen 2012). For instance, sticklebacks from lake and river in northern Germany have been reciprocally exposed to either habitat, resulting in increased parasite abundance in fish exposed to lake conditions (Scharsack *et al.* 2007). Further examples have been reported in cichlids in East African lakes (Maan *et al.* 2008) Whitefish in Swiss pre-alpine lakes (Karvonen *et al.* 2013b) and Trinidadian guppies (Fraser & Neff 2009).

Despite host genetic factors also the environment itself alters the strength as well as the response to selection during host-parasite-interactions. Temperature has been shown to play an essential role on infection rates and virulence. MacNab & Barber (2012) infected threespine sticklebacks with the tapeworm *Schistocephalus solidus* and exposed the fish to different temperatures. They could demonstrate that the tapeworm larvae grew faster at higher temperature predicting an increased output of infective parasites. Further, infected fish preferred warmer water temperatures indicating that the parasite alters their host behavior. An increased macro- and microparasite abundance and diversity at warmer water temperatures was further found by Schade *et al.* (Schade *et al.* 2015) who investigated four distinct fish species occurring in the Northern Wadden Sea.

Immigrant inviability or selection against migrants can be mediated by parasites. When an individual invades a new environment it might not be adapted to the local parasites and thus suffers from increased infectivity compared to the remote population. Moreover, organisms can escape from their common parasites by migrating to new environments lack-

ing the parasite. However, if this scenario actually happens has been questioned (Hawkes *et al.* 2007). In a transplant experiment MacColl & Chapmann (2010) demonstrated that marine sticklebacks are more susceptible to some freshwater parasite when transplanted to freshwater than the native freshwater ones and remain smaller. Also hybrid inviability can be induced by parasites; they may suffer from reduced resistance in either habitat due to their intermediate phenotype. Hence, considering that parasites and pathogens may act as an important selective force is reasonable.

Bacteria induced reproductive isolation

Bacteria influence their host

According to Brucker & Bordenstein (2012), already in the late 1920s the microbiologist Ivan E. Wallin suggested that the acquisition of bacterial symbionts was the major force in speciation processes. To date, it is indisputable that bacterial microbiota and symbiosis play a key role in eukaryotic life (Moran 2006; Fraune & Bosch 2010; McFall-Ngai *et al.* 2013). In fact, eukaryotic life would not have been developed as it has until today without bacteria (McFall-Ngai *et al.* 2013). Microbes help their host to digest food and provide digestible molecules and nutrients and are at least in part selected by the host's gut epithelium (Ley *et al.* 2005, 2008; Lozupone *et al.* 2013). In herbivorous animals they are involved in digestion and fermentation of plant polymers and the synthesis of vitamins (Chaucheyras-Durand & Durand 2009). Further famous examples for long-term coevolution are termites and their intestinal microbiota supporting them in the carbon and nitrogen metabolism. 90 % of the bacterial species in the termite hindgut cannot be found elsewhere and the gut bacteria between distinct termite genera are very different (Hongoh *et al.* 2005), implying a longtime co-evolution. Besides nutrition, microbes play a key role in animal development (Fraune & Bosch 2010). In the squid *Euprymna scolopes*, for instance, the luminescent bacterium *Vibrio fischeri* induces programmed cell death in the squid's surface epithelial tissue resulting in a light-emitting symbiotic organ wherein the bacterium stays and is supplied with peptides and proteins (Montgomery & McFall-Ngai 1994; Nyholm & McFall-Ngai 2004). The importance of bacterial microbiota in shaping its host immune system becomes clear when we consider studies about germ-free (also called gnotobiotic) animals. It has been determined that germ-free born and raised mice have defects in the development of the gut-associated lymphoid tissue and antibody production (Macpherson & Harris 2004). Moreover, they have various morphological tissue defects in the intestine (reviewed by Round & Mazmanian 2009; Hooper *et al.* 2012).

Host genetic influence on bacterial communities

Bacterial communities are on their part shaped by host genetics, physiology and lifestyle and also from environmental factors, like pH, salinity, temperature, nutrients and competitors. Although microbes are very diverse across body sites (Costello *et al.* 2009; Spor *et al.*

2011) and individuals (The Human Microbiome Project Consortium 2012), family members often have more similar microbiota than nonrelated individuals (Turnbaugh *et al.* 2009; Anderson *et al.* 2012; Yatsunenکو *et al.* 2012; Goodrich *et al.* 2014). However, besides being determined by host genetics it has been assumed that the similarity arises from being exposed to similar environments and having the same lifestyle. Indeed, household members (Song *et al.* 2013) or mice raised in the same cage (Alexander *et al.* 2006; Hildebrand *et al.* 2013) have a great accordance in their bacterial communities. Benson *et al.* (2010) published among the first study using Quantitative trait loci (QTL) analysis as tool to investigate host genetic effects on bacterial abundance and diversity. QTL enables to map associations between segregating molecular markers and quantitative traits. They compared >600 mouse advanced intercross lines and identified several host quantitative trait loci that are linked with the relative abundance of certain bacterial taxa. Additional studies using genome-wide mapping in mice (Srinivas *et al.* 2013; Wang *et al.* 2015) and fish (Boutin *et al.* 2014) followed and all found significant genetic loci in the host contributing to bacterial community structure. All these effects suggest that a specific bacterial composition is beneficial for the host; therefore host control over its microbiota is most likely.

Environmental influence on bacterial communities

Environmental or geographic effects on bacterial microbiota have been identified in several organisms, like humans (Spor *et al.* 2011), mice (Linnenbrink *et al.* 2013), fish (Sullam *et al.* 2012), insects (Yun *et al.* 2014) and amphibians (Kueneman *et al.* 2014). It has been shown, that for instance salinity (Lozupone & Knight 2007; Sullam *et al.* 2012) and pH (Walker *et al.* 2005; Duncan *et al.* 2009) significantly affect the abundance and diversity of microbes in general and in the host gut, as well as host diet (Ley *et al.* 2005; David *et al.* 2014). Among other influencing factors, like humidity and intra- or interspecific interactions between microbial species, temperature is well known to account for differences in microbial communities in several organisms (Erwin *et al.* 2012; Lokmer & Wegner 2015; Neuman *et al.* 2016). Higher temperature can increase abundance or virulence of certain bacterial taxa (Kimes *et al.* 2012) and thus can induce disease-outbreaks. The microbiota of pacific oysters (*Crassostrea gigas*) was significantly affected by temperature shifts and – stress (Lokmer & Wegner 2015). Further, the mortality rate of oysters was increased at higher temperature. A key role in microbial community structure is played by the mother's environment and the type of delivery. Human fetuses are sterile until birth and their initial bacterial microbiota resembles their mother's vaginal microbiota, if they were born vaginally, or is similar to those found on skin surface when delivered by Cesarean section (Penders *et al.* 2006; Dominguez-Bello *et al.* 2010). In their early development fish egg surfaces and later larvae are colonized by bacteria which are vertically transmitted by adults or horizontally from the aquaculture environment (Bakke *et al.* 1990; Jensen *et al.* 2004; Fjellheim *et al.* 2012; Wang *et al.* 2015).

Despite these well-known effects, microbes have been neglected in speciation research for a long time and only since the last two decades few examples for symbiont-induced reproductive isolation exist (Brucker & Bordenstein 2012).

Bacteria in fish

The bacterial microbiota of humans has been comparably well studied within the last decade (Ley *et al.* 2005; Grice *et al.* 2009; Costello *et al.* 2009; Grice & Segre 2011; Yatsunenko *et al.* 2012) but the microbiota of animals gained less attention. To date, there are 33,200 different fish species on earth (www.fishbase.org, version 1/2016) which is about half of all vertebrate species (IUCN Red List, version 9/2015) but their bacterial communities are underexplored. However, with increasing human population and resulting decline of wild fish stocks there's need in fish as food and the aquaculture industry experiences a vast growth (Joint Food and Agriculture Organization of the United Nations (FAO), www.fao.org). Accordingly, there's growing interest in replacing, at least in part, the application of antibiotics with probiotics, i.e. 'live microorganisms which when administered in adequate amounts confer a health benefit on the host' (FAO/World Health Organization (WHO) 2001) (reviewed by Balcazar *et al.* 2006). As in other organisms (Warnecke *et al.* 2007; Ley *et al.* 2008; Turnbaugh *et al.* 2009; Spor *et al.* 2011; Yatsunenko *et al.* 2012; Lozupone *et al.* 2012; Koch *et al.* 2012; Goodrich *et al.* 2014), also in fish most of the studies about microbiota concentrated on the gut (Rawls *et al.* 2004; Roeselers *et al.* 2011; Larsen *et al.* 2014a; Givens *et al.* 2015). However, the skin of an organism acts as first physical barrier against pathogens or toxic substances and is colonized by a variety of resident and mutualistic/commensal microorganisms protecting their host (Cogen *et al.* 2008; Grice & Segre 2011). In aquatic animals the skin is in direct contact with the environment and its changing abiotic and biotic factors and thus a variety of bacterial communities can be expected. In teleost fish skin lacks keratinization and secretes mucus which is important for immune functions (Salinas *et al.* 2011). Additionally, it may provide local and distinct niches for commensals, mutualists and pathogens, depending on disparate epidermal mucus secretion, variable exposure to nutrient excretion fluxes through gills and vent and variable water flow during swimming (Esteban 2012). Besides the body surface mucus can also be found on the gills and in the gastrointestinal tract and it acts as physical, biochemical and semipermeable barrier enabling exchange of nutrients, water, gases, odorants, hormones and gametes (Esteban 2012). It is also involved in osmoregulation, ion regulation, lubrication and parental care and inhibits the invasion and proliferation of pathogens (Esteban 2012). As a first defense mechanism, pathogens can be trapped by the mucus layer, for instance by decreased viscoelasticity, before entering the epithelial surface and are then removed by the surrounding water (Benhamed *et al.* 2014). Moreover, mucus is constantly renewed and digested which makes it difficult for pathogens to enter the host. Teleost fish exhibit a skin-associated adaptive immune system and contain the immunoglobulins IgM, IgD and IgT/Z (Salinas *et al.* 2011). Very little is known about the skin-associated lymphoid tissue (SALT) in fish, though it has been demonstrated to contain secretory cells producing mucus, lymphocytes, granulocytes, macrophages and Langerhans-like cells (Salinas *et al.* 2011). Langerhans cells are dendritic cells (antigen-presenting cells) of the skin and mucus, whereas lymphocytes, granulocytes and macrophages are certain types of white blood cells and are all involved in the innate immune response (Janeway 2005). Despite pathogens, the mucus of a healthy fish contains great amounts of mutualists and commensals

(Cahill 1990; Austin 2006), so there's need to find the balance between fighting against pathogens while tolerating commensals (Gomez *et al.* 2013). It has been suggested that specialized receptors of the innate immune system, e.g. transmembrane Toll-like receptors (TLRs) are able to discriminate between 'friend or foe' (Galindo-Villegas *et al.* 2012). In mammals, TLRs can recognize microbe-associated molecular patterns (MAMPs) on the pathogens and full-length orthologues have been identified in zebrafish, as well (Phelan *et al.* 2005). Mutualists/commensals in the mucus are important for the fish. They can for instance reduce drag in response to turbulent flow because they are hydrophobic and produce a smoother surface at the mucus-water interface (Sar & Rosenberg 1987; Bernadsky *et al.* 1993).

Recent studies about skin bacteria in several distinct fish species revealed more or less consistent results. Salinity has been identified to have general impact on bacterial abundance (Lozupone & Knight 2007) so it was not surprising that there are major differences in microbial communities between freshwater and saltwater fish (Roeselers *et al.* 2011; Sul-lam *et al.* 2012). Generally, fish skin harbors bacteria belonging to the major phylum Proteobacteria, followed by Bacteroidetes and Firmicutes (Larsen *et al.* 2014a). Fish from freshwater contain, amongst others, *Acinetobacter*, *Aeromonas* and *Pseudomonas*, whereas fish from saltwater often contain *Vibrio*, *Shewanella* and *Clostridium* (Larsen *et al.* 2014a). It has been hypothesized that the skin of aquatic organisms reflects at least in part, the bacteria of the surrounding water (Austin 2006). At the same time mucus-associated bacteria can be very distinct from the environmental as it has been determined for amphibians (Walke *et al.* 2014) and fish (Wang *et al.* 2010; Chiarello *et al.* 2015).

Generally, host inter-individual differences in bacterial composition are related to host physiology, age, sex, health, the immune system, as well as to environmental factors. However, it is sometimes difficult if not impossible to disentangle cause and consequence of changes in skin microbiota.

Hybrid zones as natural laboratories to study early adaptive speciation

Hybridization is, according to Barton & Hewitt (Barton & Hewitt 1985), the reproduction between members of genetically distinct populations producing offspring of mixed ancestry. It is relatively common in multicellular plants and animals: about 10-30 % per species regularly hybridize and in sympatric populations 0.01-1 % of all individuals are hybrids (Mallet 2005). Hybridization plays a clear role in speciation when the hybrids' chromosome numbers doubled (allopolyploidie) because there's instantly a chromosomal mismatch and thus reproductive isolation towards either parent (Mallet 2007). Almost 1/3 of all plants have emerged by this process (Wood *et al.* 2009). During homoploid hybrid speciation the distinct populations meet, hybridize and merge their gene pool which often lead to the emergence of hybrid zones or hybrid swarms and potentially new hybrid lineages can evolve (Nolte & Tautz 2010). In contrast to allopolyploid hybrid speciation, homoploid hybrid speciation is less obvious and more difficult to detect, however it is more common as

previously assumed (Seehausen 2004; Mallet 2007, 2008; Nolte & Tautz 2010; Abbott *et al.* 2013). In a segregating hybrid population the hybrids can exceed their parents' phenotypes, which is defined as 'transgressive segregation' (Rieseberg *et al.* 1999b). This extreme appearance can result in higher or lower fitness for the hybrids compared to their parents. Probably the most famous and well-studied examples are the hybridizing sunflower species *Helianthus annuus* and *H. petiolaris*. Three hybrid lineages arose out of this hybridizing event and all of them showed higher tolerance to drought. Hence, they colonized deserts, areas where their parents are absent (Rieseberg *et al.* 2003a; Gross & Rieseberg 2005). The outcome of a hybridization event can lead to genetic novelty by new gene combinations but can also result in the breakdown of co-evolved gene compositions and cause reduced fitness (Burke & Arnold 2001). Hybrids can also have reduced fitness due to negative epistasis between immune genes from distinct individuals, underdominance, segregation distortion and chromosome rearrangements (Nolte & Tautz 2010).

Hybridization between distinct populations often emerges at ecotones, where two habitats come together, each favoring different alleles or allele combinations. A hybrid zone arises, often with steep ecological clines within environmental gradients. If the zone width is shaped by the balance between dispersal and selection against hybrids, either ecological or genetic or both, it is called a tension zone (Hewitt 1988). Hybrids can only establish as a new lineage when becoming spatially or genetically separated from the parents (Arnold 1992; Buerkle *et al.* 2000; Seehausen 2004; Mallet 2007). Hybrid zones contain recombinants of up to thousands of generation and thus provide a promising scenario to study divergence and adaptive speciation *in situ* (Hewitt 1988). In fact, they are comparable to artificial crosses, e.g. for QTL mapping but are directly influenced and selected by nature. Thus, they can serve to identify genomic regions restricting gene flow in natural environments (Rieseberg *et al.* 1999a). Further, they enable to map species that would not be suitable for crossing experiments because of their long generation time or large physical size (Gompert & Buerkle 2009).

The *Cottus* System

European sculpins (*Cottus sp.*, Scorpaeniformes, Teleostei) are fresh water fishes that are distributed from the Baltic Sea to southern Europe and can be differentiated to 15 distinct species, based on morphological and molecular data (Freyhof *et al.* 2005; Kottelat & Freyhof 2007). They belong to the family Cottidae which contains around 100 species and is among the fastest 10 % of all ray-finned fish (Actinopterygii) concerning the speciation rates (Rabosky *et al.* 2013; Seehausen & Wagner 2014). European sculpins inhabit small, cold and oxygen-rich tributaries and are usually absent in main rivers. Surprisingly, around 30 years ago, they have been commonly found in the Lower Rhine in the Netherlands and in the Rhine Delta, in large and comparably warm and less oxygenated water bodies (De Nie 1997).

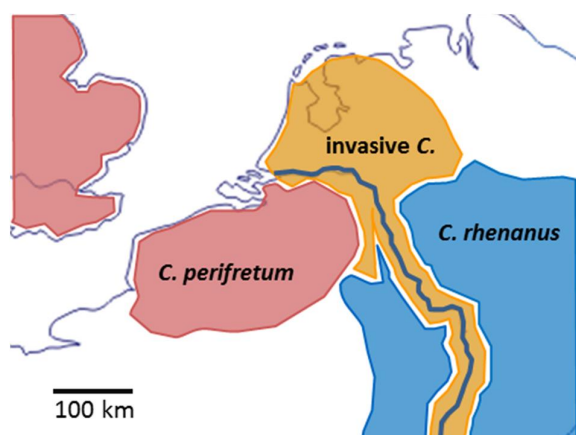


Figure 1.1: Distribution of *Cottus* spp. in Western Europe around the river Rhine basin. *C. perifretum* is inhabiting river systems in Great Britain, Belgium and the Netherlands and *C. rhenanus* is mainly occurring in the river system Rhine in Germany. Approx. 200 years ago these river systems were artificially connected; the *Cottus* species met and generated fertile hybrids. These hybrids established as a new lineage and invaded new habitats which were previously free of *Cottus*. The invasive *C.* expanded their range and dispersed upstream the river Rhine. Map modified after: <http://www.d-maps.com>

Apparently, the sculpins invaded new habitats that have been free of *Cottus* sp. by then. In the early 1990s they have been reported in the river Sieg, a major tributary of the Rhine in Germany, and expanded on average 4-8 km per year (Nolte *et al.* 2005). In-depth analysis of *Cottus* sp. from Belgium, the Netherlands and Germany determined that the invasive sculpins differed morphologically and ecologically from the resident sculpins inhabiting the small streams (Nolte *et al.* 2005). Moreover, single nucleotide polymorphisms (SNP) and mitochondrial haplotype sequence analyses revealed that the invasive fish are hybrids emerged between Rhine sculpins, distributed in lower and middle Rhine and Meuse drainages (*Cottus rhenanus*) and Scheldt sculpins, distributed in Great Britain, river Scheldt in France and Belgium, Rhine and Seine (*Cottus perifretum*) (Kottelat & Freyhof 2007). They form a homogenous population being distinct from the other groups and contain almost intermediate proportions of either parental gene pool (Nolte *et al.* 2005). Meanwhile, mitochondrial haplotype analysis revealed that at least three differentiated hybrid lineages have emerged out of this hybrid situation and additional modeling approaches suggested that there is constant gene flow from the parental species to the hybrid lineage but they nevertheless remain distinct (Stemshorn *et al.* 2011). In life history traits, like maturation, and also phenotypically the invasives rather resemble *C. perifretum* than *C. rhenanus*, for instance in the amount of prickles (Nolte *et al.* 2005). Prickles are spine-like scales covering the body. Koli (1969) and Nolte (Nolte *et al.* 2005) defined distinguishable classes from 0 = absent to 4 = prickling extends the back beyond the middle of the second dorsal fin. Presence, absence and abundance of skin prickling is a diagnostic pattern by which several *Cottus* species can be distinguished (Kottelat & Freyhof 2007). Invasive sculpins usually are, like Scheldt sculpins, intensively prickled, whereas Rhine sculpins totally lack prickles. Cheng *et al.* (Cheng *et al.* 2015) recently confirmed that skin prickling in sculpins is associated with ectodysplasin signaling genes which are known to influence ectodermal struc-

tures in human, fish and other vertebrates (Sadler *et al.* 2014).

Approximately 200 years ago man-made canals connected the Scheldt and the Rhine system and enabled the hybridization of *C. rhenanus* and *C. perflretum* which is hypothesized as the initial time point of the emergence of invasive sculpins (Nolte *et al.* 2005) (Figure 1.1). Since invasive *Cottus* expanded their range with > 4 km per year (Nolte *et al.* 2005) several new contact zones arose where small tributaries, inhabited by Rhine sculpins, disemboque into the larger rivers Sieg and Mosel during the last 20-25 years. Nolte *et al.* (2006) studied several of these secondary contact zones and with the help of microsatellite markers they determined steep sigmoidal clines with only about 2 km width within the hybrid zones. Apparently, the invasive sculpins stopped their dispersal and selection acts against pure and hybrid immigrants in both habitats. No genetic incompatibilities and hybrid inviability have been identified so that the authors hypothesized ecological-driven selection defined by an ecotone model. Population genetic analyses revealed a large number of backcross hybrids in the contact zone (Nolte *et al.* 2006); Schroeder (2009), unpubl. Master Thesis) so apparently they are viable but restricted by environmental selection. However, how environmental factors are shaping the contact zone structure has not been studied, yet. Rhine sculpins only occur in small tributaries even when invasive sculpins are absent (Nolte *et al.* 2005) whereas invasive sculpins indeed mainly colonize larger rivers but also inhabit streams in the absence of Rhine sculpins (Nolte *et al.* 2006).

SCOPE OF THE THESIS

Overall, this thesis aims to investigate whether bacterial microbiota and parasites influence adaptive speciation of European freshwater sculpins, hybridizing at an ecotone, by limiting their dispersal to foreign habitats. No intrinsic barriers between the sculpin lineages have been detected, thus it is assumed that ecological factors, for instance bacterial microbiota and parasites, are inducing selection pressure against migrants and are shaping this contact zone.

In Chapter I, I first analyzed skin bacterial microbiota of wild Rhine sculpins, invasive sculpins and their hybrids, all occurring in two habitats stream and river. I revealed that bacterial communities are mainly influenced by host habitat but also by host genotype. Second, by studying distinct lab raised sculpins exposed to the same husbandry conditions I identified some host genetic effects on bacterial composition and abundance.

Because in the sculpin system host habitat and host genetic factors were not completely independent from each other, it was necessary to examine other teleost fish occurring in both habitats stream and river in high abundance. This gave rise to Chapter II, wherein I examined the skin bacterial communities of *P. phoxinus* and *B. barbatula* and compared them to the *Cottus* bacterial microbiota. I identified similar results to those in chapter I, however, here the bacterial community structure of the fish was mainly shaped by host genotype: the bacterial composition of distinct fish species was less similar than that of the same fish species in distinct habitats. Further, I revealed that sculpins caught in 2005 and 2009 show similar patterns of bacterial composition which are independent from sampling year and method.

In Chapter III I described the results of three pilot studies investigating the parasite diversity and abundance among different sculpin lineages. First I presented a parasite screen conducted at distinct wild caught sculpins. I found several different macroparasites in the fish, whose richness and abundance varied with host genotype and was further correlated with host body size. Second, we carried out two infection experiments wherein we exposed distinct sculpin lineages to a common fish parasite and thereafter recorded the parasitic load per host. I identified that the parasite abundance varied between sampling populations but not genotype.

CHAPTER 1

Skin bacterial microbiota associated with naturally hybridizing *Cottus*

Introduction

Finding the genes underlying speciation has been of great interest for several decades (Coyne & Orr 2004; Noor & Feder 2006; Wolf *et al.* 2010; Nosil & Schluter 2011; Santini *et al.* 2012). How do species diverge? What are the forces acting on speciation processes? Which and how many genes are involved? How do barriers to gene flow evolve? These are just some of the arising questions. When studying speciation, there is the necessity to first define what a species is. Several different concepts have been formulated (De Queiroz 2007) but in accordance with Mallet (1995, 2007) I here refer to species as distinguishable groups of individuals having a shared gene pool that remains over time despite possible gene flow, indicating that barriers to reproduction exist. There are three different types of reproductive isolation: i) extrinsic postzygotic isolation, which evolves through divergent selection (ecological or sexual), ii) intrinsic postzygotic isolation, caused by genetic incompatibilities and iii) prezygotic isolation, like mechanical, behavioral and temporal barriers (Seehausen *et al.* 2014). Lately, especially the ecological impact on the evolution of reproductive isolation has gained more and more attention (Schluter 2001, 2009; Rundle & Nosil 2005; Hendry *et al.* 2007; Nosil 2012). Schluter (2009) suggested to group speciation processes into two categories: ecological speciation and mutation-order speciation. In ecological speciation, selection acts on genotypes adapted to different environments and barriers to gene flow evolve. The environment consists of abiotic factors, such as temperature, light and pH, and biotic factors like predators, pathogens, mutualists and inter- or intraspecific competitors. Different alleles may be fixed being advantageous in one environment but not in the other, so all these factors can cause adaptive speciation processes leading into divergent directions. Popular examples for ecological speciation are threespine sticklebacks (Hatfield & Schluter 1999) lake whitefish (Lu & Bernatchez 1999) and *Heliconius* butterflies (Mallet & Barton 1989). By contrast, during mutation-order speciation reproductive isolation evolves in the same environment by random occurrence and fixation of different alleles (Schluter 2009; Nosil & Flaxman 2011) but I here will not focus on this process.

Although it is well accepted that speciation caused by divergent adaptation exist, it is still unclear how barriers to gene flow evolve and which factors are driving these processes. There's a great variety of methods to understand the genetic basis underlying these barriers (Seehausen *et al.* 2014), from classical mapping techniques, to identification of the physical occurrence of single loci in the genome, to modern approaches, like whole-genome assemblies, enabled by high-throughput sequencing. However, they often require controlled crosses between parental lineages which are limited to several generations and

hence time-consuming if not impossible for many 'higher' taxa. Therefore, most studies concentrated on model organisms or domesticated animals and plants (e.g. Long *et al.* 1995). By contrast, natural hybrid zones provide the ideal setting to overcome some of these difficulties and can be used as natural laboratories (Harrison 1993). Therein, genetically distinct individuals meet, hybridize and produce admixed and fertile offspring of hundreds or even thousands of generations of recombinants. Moreover, to study the association of traits underlying speciation processes it is indispensable to do that in nature with natural influences and selective forces. Apart from being an origin of genetic variation, through hybridization transgressive segregation can lead to extreme phenotypes in hybrids, as described for some sunflower species *Helianthus* spp. (Rieseberg *et al.* 1996, 1999a; Lexer *et al.* 2003). Compared to the parents the three emerged hybrid sunflower lineages show higher tolerance for salt and drought. Generally, hybrid zones occur in a wide variety of organisms (Barton & Hewitt 1985; Mallet 2005, 2007), 25 % of all plant and about 10 % of all animal species hybridize. Mostly, they are narrow regions and often emerge at habitat clines (Barton & Hewitt 1985) where distinct environments come together.

As stated above, the environment consists of various factors whereof the bacterial microbiota certainly is highly influential but nevertheless under explored. In fact, almost all eukaryotes are associated with a high variety of symbiotic bacteria (Bevins & Salzman 2011), which influence their host's nutrition, development and immune system, and whose composition on their part may depend on host physiology, diet and immune system (Zoetendal *et al.* 2001; Rawls *et al.* 2006; Fraune & Bosch 2007; Turnbaugh *et al.* 2009) as well as on the environment (Spor *et al.* 2011; Sullam *et al.* 2012; Linnenbrink *et al.* 2013). Zebra fish, *Danio rerio*, seem to have a core gut microbiome, whether or not fish were domesticated and stem from laboratory facilities or were caught from the wild (Roeselers *et al.* 2011). Further, Franzenburg *et al.* (2013) revealed for different species of the cnidarian *Hydra*, that host-associated bacteria are determined by species-specific antimicrobial peptides. On the other hand, Sullam *et al.* (2012) indicated that the composition of fish gut bacteria differs with salinity level, indicating that the environment can have a large impact, either. Several studies proved that symbiotic microbes can facilitate reproductive isolation and finally the emergence of new species (Engelstädter & Hurst 2009; Brucker & Bordenstein 2012).

Here, I use a recently emerged secondary contact zone of two European freshwater sculpins (*Cottus* spp.) to analyze whether bacteria act as selective force on early adaptive speciation processes. Two fish lineages hybridize under absence of intrinsic barriers (Cheng *et al.* 2013) but their hybrid zone is narrow and selection acts against immigrants to foreign habitats (Nolte *et al.* 2006). Hence, extrinsic ecological factors, like bacterial symbionts, are supposed to stabilize this scenario through selection pressure.

Briefly, sculpins used to be adapted to cold and oxygen rich narrow streams, thus, it was rather surprising that since the 1980s they have constantly been found in the river Rhine and its major tributaries being comparably warm and less oxygenated (Nolte *et al.* 2005). Meanwhile, genetic analyses revealed that these invasive sculpins are hybrids between

ancient sculpin species inhabiting the river system Scheldt (*C. perifretum*) and the river system Rhine (*C. rhenanus*), which have been separated 1-2 million years ago (Englbrecht *et al.* 2000; Volckaert *et al.* 2002). About 200 years ago their habitats have artificially been connected through man-made canals, hence, this time point is assumed to be the cradle of invasive sculpins. The invasive sculpins disperse upstream in the Rhine system between 4-8 km per year (Nolte *et al.* 2005), which led to the emergence of several secondary contact zones with one of their parental lineages, *C. rhenanus* in the mid-1990s (Nolte *et al.* 2006). Ancestry coefficients plotted against river distance indicated that these zones are just about 2 km wide and contain pure ancestral individuals above and below both ends and admixed individuals in between (Nolte *et al.* 2006). Selection acts against immigrating genotypes and each lineage remains in its primary habitat. There's general evidence for environmental selection rather than for genetic incompatibilities (Nolte *et al.* 2006, 2009).

In this study I focus on the secondary contact zone where the small stream Broel, inhabited by Rhine sculpins, disembogues into the larger river Sieg, inhabited by invasive sculpins, near Bonn, Germany (Figure 2.1). To analyze to what extent host-bacterial composition is shaped by host genetic architecture I study the genomic ancestry of individual fish and determine associations with bacteria living on their skin. Only few studies concentrated on microbiota on the skin of fish, although its direct exposure with the surrounding water and biotic and abiotic factors makes it an interesting study object. In fish, epidermis and mucus serves as first physical barrier of the immune system (Esteban 2012; Gomez *et al.* 2013; Benhamed *et al.* 2014), preventing pathogen colonization and infection and thus has an important role to discriminate between pathogenic or commensal bacteria. Further, focusing on fin skin enables to take small tissue samples without sacrificing the whole animal, which is in agreement with the recent revise of the legislation regarding animal experiments (Directive 2010/63/EU). Subsequently, it enables to study endangered organisms in which sampling is prohibited by law (Czypionka *et al.* 2015). As described above, the respective contact zone contains pure and recombinant individuals so that I am able to investigate three different genetic groups: Rhine sculpins, invasive sculpins and their hybrids, all occurring in both habitats stream and river.

First, I will characterize the bacterial communities of sculpins and further disentangle to what extent host genetic and environmental effects affect their microbial composition. Thus I can examine whether adaptation to certain associated bacteria acts as natural selection against immigrant genotypes. Second, by controlled laboratory experiments I will be able to investigate host genetic effects at minimal environmental variation. Since microbiota is also shaped by host genetic architecture I expect to find differences between distinct fish lineages. Third, I will show whether bacterial communities in wild caught and lab-reared sculpins have some accordance. If every host has a specific microbiota, influenced by its lifestyle, phylogeny and genetic architecture, which is shaped independent from the surrounding, this would predict similarities in bacterial composition between hosts from the wild and from the lab. At the same time I expect to find large differences between wild caught and lab-reared fish since have been exposed to very different environmental factors.

Methods

Study populations and tissue sampling

Wild caught fish from the contact zone Sieg/Broel

Fish from a natural hybrid zone were studied for their bacteria to analyze whether certain *Cottus* spp. genotypes are associated with different bacteria in different environments. The fin samples used here stem from a large sampling in 2005. In total, 768 fish were sampled at 10 different sites along the contact zone Sieg/Broel, near Bonn, Germany (Figure 2.1). Sampling sites 1 -4 are located at the Sieg and in the following defined as river habitat. Site 5 is the hybrid zone, since here the stream Broel disembogues into the river Sieg, and sites 6-10 are at the Broel and defined as habitat stream. Fish were stunned with portable electro shockers, a third of the right fin was cut and transferred to a tube containing 500 μ l HOM-buffer (80 mM EDTA, 100 mM Tris, 0.5 % SDS) and 5 μ l Proteinase K (NEB 20 mg/ml).

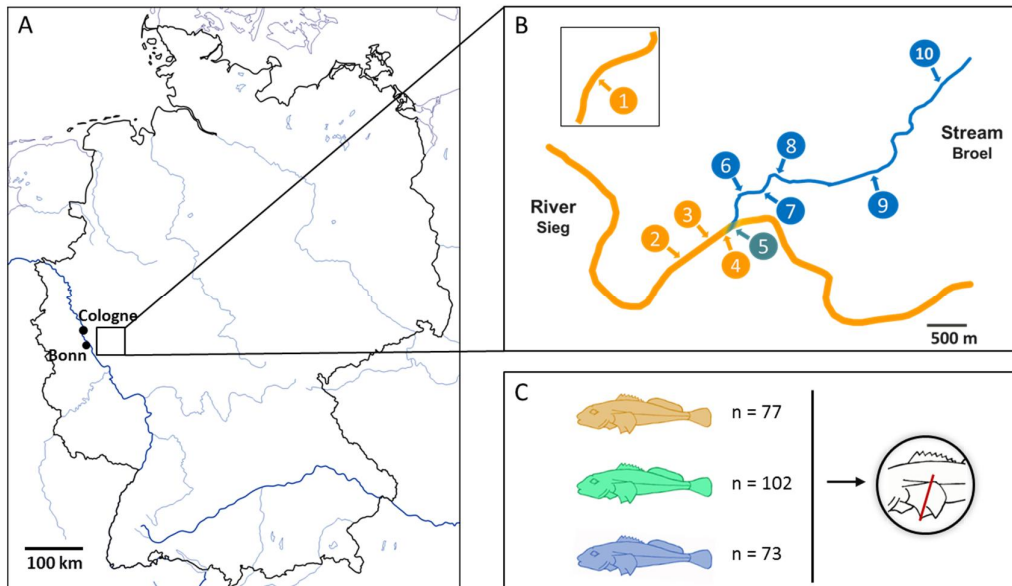


Figure 2.1: Sampling location of wild caught sculpins and experimental design. (A) The contact zone is near Bonn, Germany. (B) Fish were collected in river Sieg (site 1-4), the contact zone (site 5) and stream Broel (site 6-10). (C) They were stunned and a third of one pectoral fin was taken for microbial and SNP analysis; the latter revealed a genotype distribution of invasive *Cottus*: n = 77 (orange) hybrids: n = 102 (green) and *C. rhenanus*: n = 73 (blue). Map modified from d-maps.com.

Table 2.1: Origin of fin samples with reference number, sampling site description, GPS data, number of samples and habitat classification. The classification is based on a study of Nolte *et al.* in 2006.

No.	Site	Northern Latitude	Eastern Longitude	Classification
1	River Sieg at Meindorf	50°46'55.50"	7°7'47.78"	River
2	River Sieg at Allner, between bridges	50°46'40.32"	7°17'56.65"	River
3	River Sieg at Allner, bridge	50°46'49.56"	7°18'17.39"	River
4	River Sieg at Allner	50°46'50.07"	7°18'19.69"	River
5	Mouth of Broel south-west of Hennef	50°46'53.50"	7°18'27.09"	Hybrid Zone
6	Stream Broel at Mueschmuehle, between bridges	50°47'3.74"	7°18'30.55"	Stream
7	Stream Broel at Mueschmuehle, bridge	50°47'5.18"	7°18'39.98"	Stream
8	Stream Broel south-east of Broel	50°47'10.27"	7°18'53.09"	Stream
9	Stream Broel south of Broel	50°47'13.37"	7°19'46.73"	Stream
10	Stream Broel north-east of Broel	50°47'45.97"	7°20'23.24"	Stream

Laboratory bred fish

The common garden experiment aimed at investigating whether the sculpin lineages from the contact zone Sieg/Broel, *C. rhenanus* and invasive *C.*, differed in their bacterial composition when the environmental conditions are the same. Therefore, I raised offspring of wild caught parents, which spawned in the lab in February 2013. First, the young were fed with brine shrimps and later with frozen and alive insect larvae. The room temperature and light duration simulated the respective conditions in central Europe throughout the year. The actual experiment was started with 6 fish per lineage per aquarium tank and 12 replicates, resulting in 144 fish in total (Figure 2.2). Each tank contained 100 ml gravel and 10 L water and was continuously aerated. The water was manually changed every second day. Over the time we lost 80 juvenile fish due to an unknown infection, so that only 64 remained for further analysis. After 108 days the remaining fish were euthanized with an overdose MS222 (1 g/L) and immediately put in 70 % EtOH, each in a sterile, individual 15 ml Eppendorf tube. Some *Cottus* species can be phenotypically distinguished by presence, absence or abundance of spine-like scales embedded in the skin, also called prickles (Koli 1969). It has been shown before that invasive sculpins have prickles all over the body, whereas Rhine sculpins are mostly unprickled (Nolte *et al.* 2005). I used this morphological pattern to phenotypically differentiate the lineages, since the fish have been kept in mixed groups and couldn't be labelled. After 3 h in 70 % EtOH the skin was dried out and hence possible prickles could be detected easily with a microscope. Used instruments were flamed between the fish to avoid cross contaminations. Finally, fish were stored in sterile 15 ml Eppendorf tubes filled with RNAlater and frozen at -20 °C until further analysis.

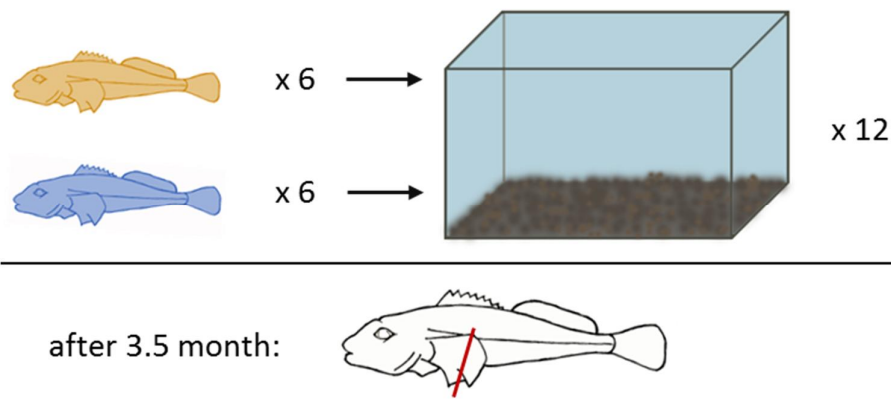


Figure 2.2: Design of common garden experiment. Six fish per species, invasive *C.* (orange) and *C. rhenanus* (blue), were set together in one tank, replicated by 12 and resulting in 144 fish in total. After 3.5 month fish were sacrificed and a third of one pectoral fin per fish was taken for microbial analysis.

Sample processing for SNP analysis

The fin samples were originally sampled to analyze the genetic structure across the hybrid zone. Within the scope of a master thesis (2009, unpubl.), Schroeder typed the sculpins for 85 single nucleotide polymorphism (SNP) markers (Stemshorn 2007) to assess their genetic background. Briefly, DNA has been extracted from fins stored in HOM buffer using a standard Proteinase K/Chloroform protocol, purified by ethanol precipitation and eluted in TE buffer. The primer-extension based method SNaPshot® Multiplex System (Life Technologies), including a Multiplex-PCR, served to type the SNP loci and data was edited with GeneMapper® (Applied Biosystems™) (Stemshorn 2007; Stemshorn *et al.* 2011); Schroeder 2009, unpubl. Master Thesis). To infer the most likely number of populations, K , based on multilocus genotype data the software STRUCTURE 2.1 (Pritchard *et al.* 2000; Falush *et al.* 2003) was used. The parameters followed the approach of Nolte *et al.* (2006): Burnin 50.000, Run Length 50.000 and 5 repeats for each value of K between 1 and 6. The SNP data enable to precisely estimate the genomic ancestry of every individual.

Sample processing for bacterial 16S rRNA gene sequencing

For the present study, I chose fish samples such that I gained a sufficient amount of all possible combinations of the three fish genotypes occurring in the three different habitats: (i) river, (ii) hybrid zone and (iii) stream. However, very few samples from hybrid zone remained after sequence processing, so I excluded them from the dataset. At the end, I performed statistical analyses with Rhine sculpins from river (rhenRiv $n = 8$), Rhine sculpins from stream (rhenStr $n = 64$), hybrids from river (hybrRiv $n = 44$), hybrids from stream (hybrStr $n = 59$), invasives from river (invRiv $n = 65$) and from stream (invStr $n = 12$).

DNA Extraction

For the microbial analysis, a third of both pectoral fins and the caudal fin of each fish were cut off with sterile instruments, rinsed with sterile 1x PBS (phosphate-buffered saline) and transferred to a 2 ml screw-cap tube containing 50 mg each of 0.1 mm, 0.5 mm and 1 mm glass beads (BioSpec Products). The bead containing tubes were autoclaved and exposed to UV light for 2 h before use. The DNA was extracted with the QIAmp® DNA stool mini kit (Quiagen) with few modifications. After pipetting 1.4 ml ASL lysis buffer to the tissue, samples underwent a bead beating using a tissue homogenizer (Precellys®) for three times 15 s at 6500 g. The suspension was heated for 10 min at 95°C and centrifuged for 1 min at maximum speed. Subsequently, the manufacturer's protocol was followed.

Library preparation

To analyze the bacterial composition of each fish I used the dual-index sequencing strategy developed by Kozich *et al.* (2013). The 16S rRNA genes were amplified with custom primers (Caporaso *et al.* 2011) spanning the V4 hypervariable region 515F (AATGATACGGCGACACCGATCTACAC*<i5>*TATGGTAATTGTGTGCCAGCMGCCGCGGTAA) and 806R (CAAGCAGAA-GACGGCAACGAGAT*<i7>*AGTCAGTCAGCCGGACTACHVGGGTWTCTAAT). The adaptor sequences which allow annealing to the flow cell of the Illumina MiSeq are indicated in bold. A 8-nt index sequence, *i5* and *i7*, respectively, enables to individually barcode each sample. This is followed by a 10-nt pad sequence to boost the melting temperature and a 2-nt link. Finally, the underlined 19- and 20-nt, respectively, represent the conserved bacterial 16S rRNA gene sequence.

I generated two libraries in which samples of all considered groups of interest were distributed as illustrated in Table 2.2.

Table 2.2: Final sample distribution in the two libraries. Shown are the respective fish genotypes per library and habitat. Some samples were removed during sequence processing so that the sample size of the groups varied. Two groups contained < 3 samples and were excluded from statistical analyses.

	Library/Run I		Library/Run II	
	River	Stream	River	Stream
<i>C. rhenanus</i>	6	19	2	46
Hybrids	26	17	18	41
invasive <i>C.</i>	40	2	25	10

Additionally, I included 15 samples in both libraries, as technical replicates, to later estimate a possible variation between the runs. Both libraries in this project were prepared equally if not stated differently. PCR was carried out using the following conditions: 98°C for 30 s, 98°C for 9 s, 55°C for 15 s, 72°C for 20 s, 72°C for 10 min with 30 cycles (library I) or 32 cycles (library II), including negative controls for each barcode combination and a positive control per PCR reaction. The mastermix contained 5 µl DNA, 13.75 µl (library I) or 12.75 µl (library II) PCR water, 5 µl buffer, 0.5 µl dNTPs, 0.5 µl (library I) or 1 µl (library II) forward and reverse primer, each, and 0.25 µl Phusion II Hot Start Polymerase. Amplicons

of the expected size (~ 300 bp) were excised from 2 % Agarose Gel and extracted by using the MiniElute® Gel Extraction Kit (Qiagen) according to the manufacturer's protocol. I had to exclude 22 samples because we couldn't get a sufficient amount of 16S rRNA for sequencing. Afterwards, the DNA concentration of each sample was quantified with Quant-iT™ dsDNA Broad-Range Assay Kit (Life Technologies) on a NanoDrop® ND-3300 Fluorospectrometer and all samples were pooled in equimolar concentrations. Then, the pool was purified with AMPure beads (Agencourt) and finally quantified with an Agilent 2100 Bioanalyzer using the Agilent DNA 7500 Kit (Agilent Technologies). The amplicon libraries were sequenced on the Illumina® MiSeq system using the V3 (library I) and V2 Kit (library II), respectively. Both were mixed with Illumina-generated phiX control libraries and contained 364 samples (library I: n = 169, library II: n = 195), in total, whereof 15, each, were technical replicates to compare the library quality.

Sequence processing

Raw sequences were de-multiplexed using Casava v.1.8.2, the paired reads were then quality filtered, trimmed and merged using USEARCH v. 8.0.1616 (Edgar 2010). From that point on I followed the analysis pipeline within the software MOTHUR v.1.31 (http://www.mothur.org/wiki/MiSeq_SOP) (Schloss 2009). Sequences were aligned to SILVA database (Pruesse *et al.* 2007; Quast *et al.* 2013) and those which didn't align to the correct positions were removed. The ends were trimmed so that all sequences overlap at the same region. After identifying the unique sequences I pre-clustered the data for further de-noising (Schloss *et al.* 2011). Chimeras were eliminated by UCHIME algorithm (Edgar *et al.* 2011), implemented in MOTHUR. A Bayesian classifier was used to classify the sequences against the Ribosomal Database Project (RDP) training set v. 9 (Wang *et al.* 2007) with a 80 % bootstrap threshold and everything classified as 'chloroplast', 'mitochondria', 'unknown', 'archaea' and 'eukaryota' was removed. I created a phylum-formatted distance matrix to cluster the sequences to operational taxonomic units (OTUs) with average neighbor algorithm at 97 % similarity level. By randomly taking the same number of sequences from each sample I rarefied the samples to 2997 reads, each, to be able to perform and compare diversity measurements at the same sequencing depth. Finally, 252 samples of the contact zone and 53 samples of the common garden study as well as 15 x 2 technical replicates remained for further analysis.

Statistical analysis

All statistical tests were performed using a rarefied subset of 2997 reads/sample and were done in the R environment (R Core Team 2015), if not stated otherwise. Since I unexpectedly had to perform analyses within the runs the design was quite unbalanced and I obtained only two samples of the invasive sculpins from the stream in run 1 and of the Rhine sculpins in the river in run 2. Therefore, I had to exclude these individuals from the analyses. α -diversity metrics measure diversity within a subunit, here bacterial diversity in a

fish sample, and there's a large number of indices which all function differently depending on how they weigh species richness, i.e. the number of species, and species evenness, i.e. species equitability (Magurran 1988). Their appropriateness depends on their performance, the respective hypothesis and the data set. Therefore, I chose to measure mere species richness (number of observed OTUs), estimated species richness (Chao 1 Index) and species diversity (Shannon Index) with tools implemented in MOTHUR. The indices were squared or square root transformed if necessary, to meet the normality requirements of the residuals, and linear models with run, genotype and habitat as main effects were applied. The best model was chosen by using Akaike's information criterion (AIC). Because I lost many fish in the common garden experiment few tanks only contained fish from one genotype and therefore I had to exclude them from the analyses. Accordingly, the second run had only two invasive sculpins left and I continued my analyses with run1, only. For the remaining seven tanks I calculated the mean per genotype per tank for each index so I applied statistics on seven invasive and eight Rhine sculpin values.

β -diversity metrics measure species diversity per habitat or population of interest and I chose measurements based on abundance (Bray-Curtis) and presence/absence (Jaccard) of shared OTUs with the 'vegan' package v. 2.3-0 in R (Oksanen *et al.* 2015) and analyzed the metrics with permutational multivariate analysis of variance (ADONIS) and principal coordinate analysis (PCoA) without constraints, implemented in vegan. A PCoA is based on eigenvalue equation but can use any dissimilarity index and distances between points in the plot reflect original distances (Zuur *et al.* 2007). Phylogenetic distance metrics (weighted and unweighted UniFrac) (Lozupone & Knight 2005; Lozupone *et al.* 2011) were obtained with FastTree v. 2.1.7 (Price *et al.* 2009, 2010) and MOTHUR and subsequently analysed with ADONIS and PCoA, as well. In the common garden experiment the factors run and genotype were nested in aquarium tank so I corrected for that when necessary. I also applied a correspondence analysis (CCA) on the OTU table which analyses unimodal relationships between species and environmental variables (Zuur *et al.* 2007). To identify associations between species, here OTUs, and combinations of sites, here fish individuals, I performed an Indicator species analyses (Dufrene & Legendre 1997; Caceres & Legendre 2009; De Caceres *et al.* 2012) with the package 'indicspecies' in R based on 10^5 permutations and p-values were corrected for multiple testing with false discovery rate (FDR) (Benjamini & Hochberg 1995). The indicator value (IndVal) index is the product of two indices:

- A (specificity) is the probability that the surveyed site (fish) belongs to the target site given the taxon was found and is based on abundance data.
- B (sensitivity) is the probability that a taxon found at the surveyed site (fish) belongs to the tested site group, it is the relative frequency.

I only considered IndVals with $A \& B > 0.5$ to exclude bacterial taxa which were only present in one fish or in few fish from the target group.

Results

The aim of this study was to characterize to what extent host genetic and environmental factors shape bacterial microbiota associated to fish skin. Briefly, after analyzing the population structure of the fish inhabiting the contact zone Sieg/Broel resulting in defining three distinguishable genetic groups, I hence used the V4 region of the 16S rDNA marker gene to generate bacterial sequence libraries of individual fish fin samples. Included were fish caught in the contact zone ($n = 252$) and fish raised in the lab under common conditions ($n = 53$). I sequenced the libraries in two distinct runs (Table 2.2) and identified that sequencing run significantly explained a major part of the variation in the data. Thus, I decided to analyze the measurements within the runs. In sum I revealed that particularly the bacterial diversity within the fish samples and the frequency of the bacterial taxa among the groups differed between the runs. However, the overall signal, indicating that environment as well as host genotype is influencing bacterial composition in sculpins, is still present.

In the following I will first present the results of the contact zone, including SNP analysis of the sculpins and subsequent characterization of bacterial taxonomy and diversity. Next, I will assess the common garden experiment and also reveal bacterial diversity within and between fish samples. Finally, I will compare the bacterial composition of wild caught and lab raised invasive *C.* and *C. rhenanus*.

Contact zone

In the contact zone Sieg/Broel two distinct sculpin populations hybridize with strong admixture

In order to investigate the population structure of fish inhabiting the secondary contact zone Sieg/Broel near Bonn, Germany, fin samples of >500 individuals have been collected in 2006. They have been genotyped and subsequent analysis with the program STRUCTURE indicated that the fish can be distinguished into two distinct populations (there was a strong increase after $K = 2$ with a tendency to plateau for $K > 2$) (Schroeder 2009, unpubl. Master Thesis). This result perfectly resembles what has been described before with morphological and ecological studies (Nolte *et al.* 2005, 2006; Stemshorn *et al.* 2011). The population structure along the contact zone reveals strong admixture between Rhine and invasive sculpins around the center which decreases with increasing distance (Figure 2.3).

The respective margins of the contact zone mainly contain pure ancestral fish but in both habitats, stream and river, also immigrant fish with pure ancestry were found. Consequently, I can compare the bacterial communities of fish with the same genotype occurring in different habitats and further of different fish genotypes occurring in the same habitat. Moreover, disentangling host genetic from environmental factors is necessary to determine whether bacteria are host-specific and thus act as a driver to early adaptive speciation processes. Based on the resulting population structure I defined three distin-

guishable genetic groups based on percentage of *C. rhenanus*: rhen (> 90 %), hybr (90 % - 10 %) and inv (< 10 %).

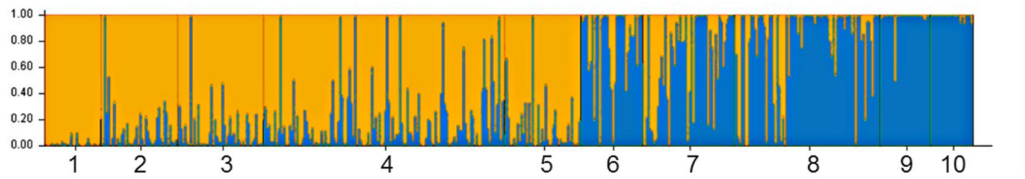


Figure 2.3: Population structure of the contact zone Broel-Sieg. Individual based genotype clustering with STRUCTURE reveals a large fraction of individuals with mixed ancestry (orange = invasive *Cotus*, blue = *C. rhenanus* ancestry). The numbers refer to the sampling sites in Figure 1B. (Plot modified after Schroeder 2009, unpubl. Master Thesis).

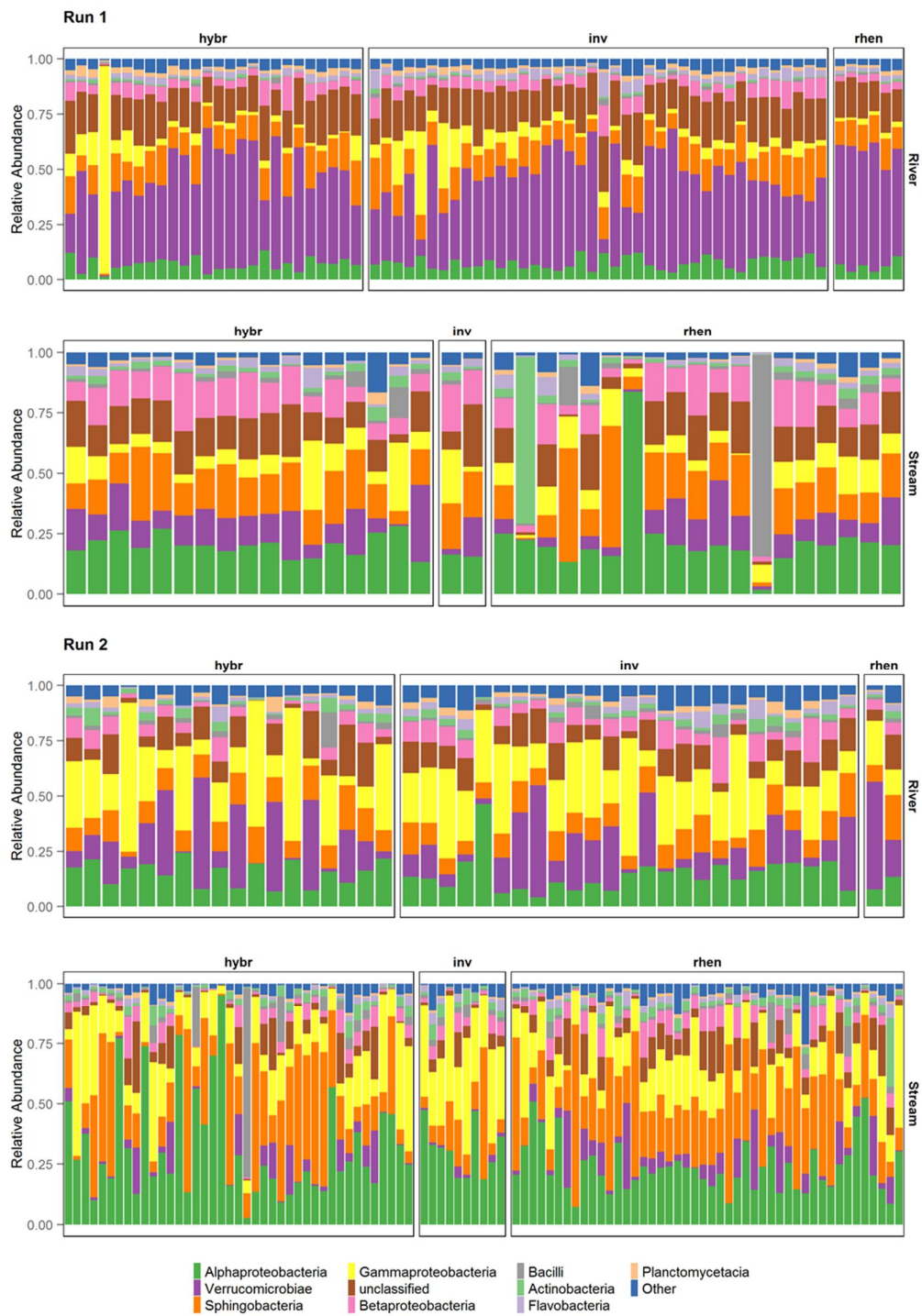


Figure 2.4: Relative frequencies of bacterial classes of *Cottus* spp. in the contact zone. Shown is the read abundance of the major classes per 2997 reads/sample. Samples are grouped according to habitat and genotype per run1 (A) and run2 (B). The taxa composition clearly differed between stream and river. Differences among the genotypes were mainly significant within the stream (see results).

Bacterial communities in sculpins are influenced by environmental factors

To characterize the bacterial microbiota in wild caught sculpins I focused on the following 252 samples caught at sampling sites 1-4 in river Sieg and 6-10 in stream Broel (Figure 1.1): Rhine sculpins from river (rhenRiv n = 8), Rhine sculpins from stream (rhenStr n = 64), hybrids from river (hybrRiv n = 44), hybrids from stream (hybrStr n = 59), invasive sculpins from river (invRiv n = 65) and from stream (invStr n = 12). I obtained only few genotypes in foreign habitats (n = 8 and n = 12), thus disentangling environmental from genetic factors in the analysis is impeded. I rarified the read number to 2997, which was the lowest number of reads per individual sample. The aim was to reveal to what extent variation in bacterial microbiota is influenced by host habitat and host genotype.

Overall, the habitats stream and river, as well as the fish genotypes within each habitat significantly differed in several major taxa. Proteobacteria was the most abundant phylum (41.2 %), followed by Bacteroidetes (23.9 %), Verrucomicrobia (20.2 %) a proportion of unclassified reads (6.1 %) and Firmicutes (3 %). Within Proteobacteria, the class Alphaproteobacteria had the highest abundance (18.7 %) besides Gamma- (14.9 %) and Betaproteobacteria (6.2%). Other major classes were Verrucomicrobiae (18.3 %), Sphingobacteria (17.8 %) and there was a proportion of unclassified reads (11.8 %). Since I obtained large differences between both runs in all major taxa, I decided to analyze them separately.

In the first run, the habitats stream and river significantly varied in all major phyla (KRUSKAL-WALLIS RANK SUM TESTS $P < 0.05$ in all cases) except for Bacteroidetes (KRUSKAL-WALLIS RANK SUM TEST, $X^2 = 3.7576$, $P = 0.053$). The fish from river had a decreased abundance of Proteobacteria (21.1 %) but a heavily increased amount of Verrucomicrobia (43.7 %). In the second run all major phyla but Firmicutes (KRUSKAL-WALLIS RANK SUM TEST, $X^2 = 2.2406$, $P = 0.134$) significantly differed between the habitats (KRUSKAL-WALLIS RANK SUM TESTS $P < 0.05$ in all cases) and the fish in the river contained much more Verrucomicrobia than those in the stream, either. At class level, the habitat comparison varied between the runs, as well (Figure 2.4). In the first run, fish from stream were enriched in the abundance of Alphaproteobacteria (KRUSKAL-WALLIS RANK SUM TEST, $X^2 = 49.96$, $P < 0.001$) and had less unclassified reads (KRUSKAL-WALLIS RANK SUM TEST, $X^2 = 4.4764$, $P = 0.034$) than fish from river. By contrast, in the second run the fish differed in all major classes besides Alphaproteobacteria.

Further, I compared the read abundance of the same fish lineages between the habitats to identify a potential habitat effect independent from host genetics. However, I obtained few parental individuals in foreign habitats and as illustrated earlier, I excluded invasive sculpins in the stream in the first and Rhine sculpins in the river in the second run due to low sample size (n = 2). In the first run I determined significant variation for Rhine (stream: n = 19, river: n = 6) and hybrid sculpins (stream: n = 17, river: n = 26) between stream and river in three and four phyla, respectively. In both lineages individuals in the stream contained an increased abundance of Proteobacteria (rhenStr: 42.4 % vs. rhenRiv: 13.3 %, hybrStr: 44.1 % vs. hybrRiv: 21.8 %). In contrast, they had less Verrucomicrobia than the fish in the river (rhenRiv: 55.5 % vs. rhenStr: 10.2 %, hybrRiv: 43.5 % vs. hybrStr: 14.3 %). At class level both Rhine and hybrid sculpins only significantly differed in the abundance of

Alphaproteobacteria which was enriched in the stream (rhenRiv: 7.3 % vs. rhenStr: 23.4 %, hybrRiv: 10.4 % vs. hybrStr: 28.7 %). In the second run invasive (stream: $n = 10$, river: $n = 25$) and hybrid sculpins (stream: $n = 41$, river: $n = 18$) significantly varied in the abundance of two and three major phyla, respectively, and mirrored the results of the first run. Besides, the hybrids showed significantly different abundances of Sphingobacteria, Betaproteobacteria and unclassified reads between stream and river (KRUSKAL-WALLIS RANK SUM TESTS $P < 0.05$ in all cases).

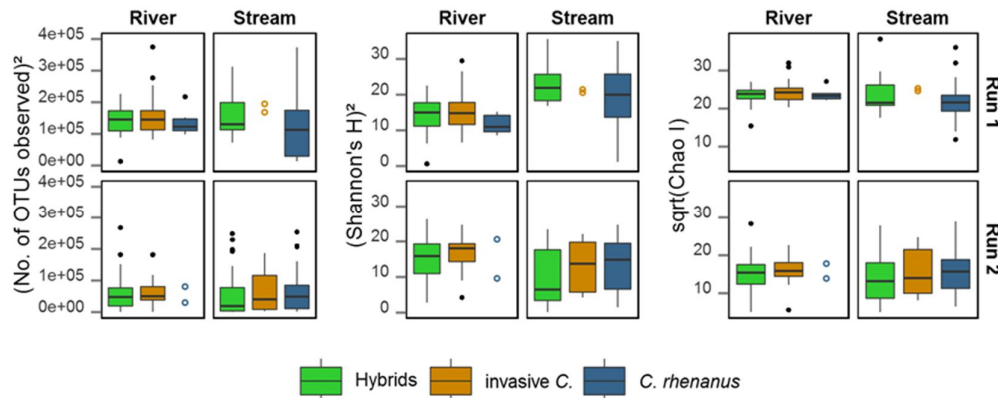


Figure 2.5: α -diversity within contact zone samples, grouped by run, habitat and genotype. Illustrated are species richness (no. of observed OTUs), diversity (Shannon's H) and estimated richness (Chao I). For groups with $n \leq 5$ no box plots just data points are shown and no statistics were applied. The richness and estimated richness was increased in run 1 and there was a significant interaction between run and habitat in the species diversity (see results).

By using α diversity metrics (species richness, diversity and estimated richness) and subsequent application of linear models I next assessed whether bacterial communities differed within individual fish. As above, I detected a notable influence of the factor run for richness and estimated richness ($P < 0.005$ in both indices). Further, in both habitats the bacterial richness in the fish was generally higher in the first run than in the second run (Figure 2.5). In addition, run and habitat were significantly interacting for the Shannon index (LM, $F_{1,248} = 32.786$, $P < 0.005$): the bacterial diversity of fish in the stream was high in run 1, whereas it was greatly decreased in run 2 and vice versa in the river. Further statistics for all models are described in Tables 2.S2, 2.S3.

Next, I evaluated the bacterial community structures between the fish (β -diversity) and first applied permutational multivariate analysis of variance (ADONIS) on Bray-Curtis and Jaccard indices, based on species abundance and presence/absence. Then, I calculated UniFrac Weighted / Unweighted indices which are phylogenetic measurements comparing the unique vs. shared branch length based on weighting taxon abundance and presence/absence, respectively (Lozupone & Knight 2008). Results for all metrics indicated that the sculpin-associated microbes are influenced by habitat, host genotype and sequencing run (ADONIS, $P \leq 0.001$ in all cases) with significant interactions between run and genotype and run and habitat (Table 2.S4). I also analyzed the metrics for both runs separately and again obtained significant effects of habitat and genotype in all metrics but UniFrac Weighted in run 2 (Table 2.S5). However, in the first run habitat always accounted for

more variation in the data than genotype but in the second run habitat was only in Bray-Curtis index the major influence (Table 2.S5). Ordination plots of PCoA reflect the results by illustrating a clear separation of habitats, particularly in run 1, and a weaker separation of genotype, rather in run 2 than in run 1 (Figure 2.8).

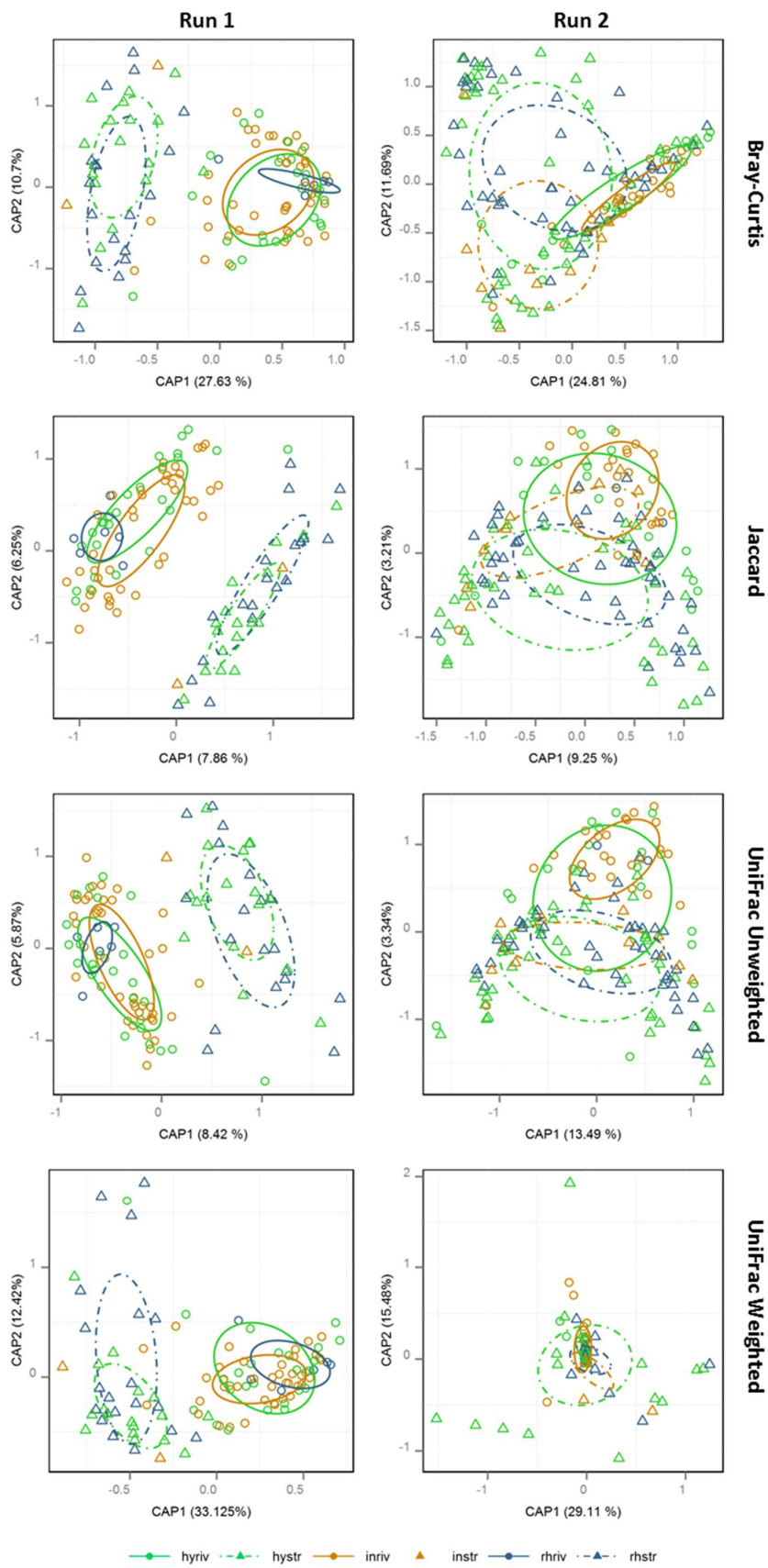


Figure 2.6: Principal coordinate analysis (PCoA) of bacterial β -diversity of fish samples from the contact zone, grouped by run. Illustrated are measurements based on OTU abundance (Bray Curtis and Weighted UniFrac) and presence/absence (Jaccard and Unweighted UniFrac). Ellipses denote the standard deviation around the mean of the respective group and were applied if the group contained $n \geq 5$ samples. In all metrics except for UniFrac Weighted in run 2 the included factors habitat and genotype are significantly affecting the bacterial composition (ADONIS $P = 0.001$). However, in run 1 habitat always accounted for most of the variation whereas in run 2 genotype accounted for slightly more variation than habitat (see Table 2.S5).

Additional to PCoA, I performed a canonical correspondence analysis (CCA), which directly uses the OTU abundance table and shows just the variation that can be explained by the available environmental factors. I performed a partial CCA with run as conditional factor to remove the unwanted effect of nestedness. I applied an ANOVA-like permutation test to evaluate the significance of the constrained axes, variables and the marginal terms of the model (Table 2.S6). I chose the model that explained most of the inertia and similarly contains least factors. Accordingly, the best model (total inertia: 15.17, explained inertia: 0.41/2.7 %, conditioned inertia: 0.31/2.1 %) included the factors habitat ($F_{1,247} = 4.72$, $X^2 = 0.28$, $P < 0.001$) and genotype ($F_{2,247} = 1.14$, $X^2 = 0.13$, $P = 0.032$) (Table 2.S6). I obtained three axes, of which one was significant (Table S6). Both runs illustrate a clear separation of the habitats stream and river and thereby reflect the patterns of the PCoA ordination (Figure 2.6). In the first run, the samples from river cluster closely together whereas the samples from stream are more dispersed, particularly the Rhine sculpins (Figure 2.7). The second run distinguishes invasive and hybrid sculpins in the river, whereas the fish from the stream overlap.

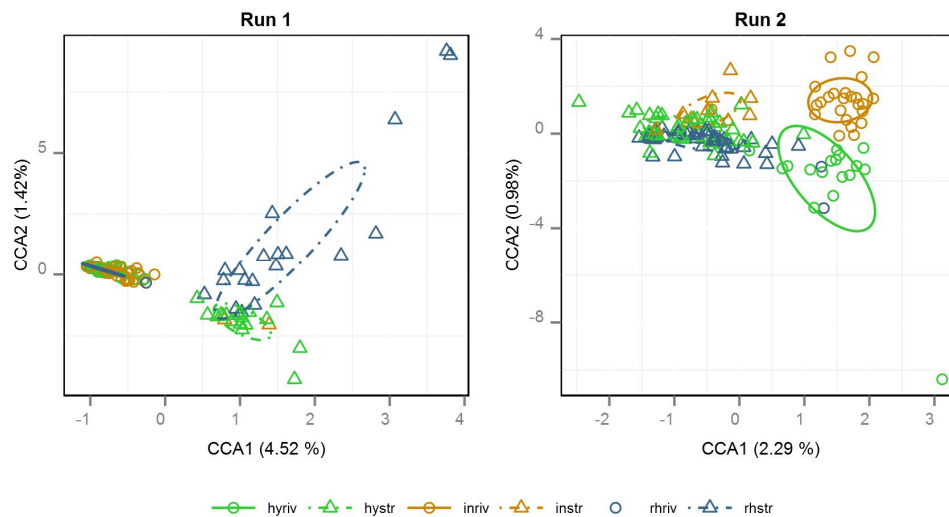


Figure 2.7: Ordination of canonical correspondence analysis (CCA) of contact zone samples conditioned by run. Plotted are the first two axes whereof the first was significant (permutation test $P < 0.001$). Included factors were habitat (stream, river) and fish genotype (hybrids, Rhine sculpins, invasive sculpins). Ellipses denote the standard deviation around the mean of the respective group and were applied if the group contained $n \geq 5$ samples. Both factors were significantly influencing the bacterial composition (permutation test $P \leq 0.032$, see Table 2.S6).

The indicator species analysis calculates species-site-associations and provides the opportunity to identify OTUs which are significantly associated to a certain group of interest, like

fish from stream or river. I analyzed both habitats and obtained in total 62 significant OTUs (river: $n = 53$, stream: $n = 9$) after correcting for multiple testing with FDR (Benjamini & Hochberg 1995). The majority of all OTUs belonged to the phylum Bacteroidetes (river: 35.8 %, stream: 11.1 %) followed by Proteobacteria (river: 18.9 %, stream: 22.2 %) and reads that could not be classified to any taxonomic level (river: 18.9 %, stream: 0.0 %). Fish in the stream mainly contained significant OTUs belonging to the phylum Actinobacteria (55.5%). At genus level the majority of all significant OTUs could not be classified (72.6 %). Besides, fish from the river mainly harbored *Flavobacterium* (3.8 %), whereof some species, like *F. psychrophylum*, are well known fish pathogens (Nematollahi *et al.* 2003, Bernadet & Bowman 2006; Starliper 2011), and some typical freshwater bacteria, like *Ferruginibacteria*, *Haliscomenobacter* and *Luteimonas*. Fish from stream mainly contained *Corynebacterium* (33.3 %), which is widely distributed in skin, e.g. in humans (Tlaskalová-Hogenová *et al.* 2004).

Different fish lineages have different bacterial composition in the same habitat

Besides habitat effects on bacterial composition I identified a significant influence of host genotype in several analyses. The frequency distribution of several major bacterial phyla and classes significantly differed between hybrids, Rhine and invasive sculpins, within stream as well as within river (Figure 2.5), but only in run 2 and not in run 1 (KRUSKAL-WALLIS RANK SUM TESTS $P < 0.05$ in all cases). However, fish in the stream were generally more diverse than those in the river. These results were supported by β -diversity analyses revealing a significant influence of genotype only in run 2 (Table 2.S5). Also the ordination plots of PCoA illustrate a clearer separation of the three sculpin groups in run 2 (Figure 2.7). Canonical correspondance analysis identified a significant influence of genotype on bacterial community structures either (ANOVA, $P = 0.032$).

Next, I applied an indicator species and analyzed each genotype (inv, rhen, hybr) per habitat (stream, river), separately. This resulted in 50 significant OTUs in total after correcting for multiple testing with FDR (Benjamini & Hochberg 1995) 46 OTUs were significantly associated to fish from the river and four to fish from the stream; the latter were shared between all three fish genotypes. Within the river Rhine sculpins contained most of all significant OTU ($n = 46$), followed by hybrids ($n = 38$) and invasive sculpins with least OTUs ($n = 36$). Overall, the majority of the significant OTUs belonged to the phylum Bacteroidetes (32 %), followed by unclassified reads (20 %), Verrucomicrobia and Proteobacteria (both 18 %). Only five OTU could be classified to a known genus and I revealed the genera *Flavobacterium*, *Luteimonas*, *Sphingomonas* and *Streptococcus*, reflecting the results of the indicator species analysis across the habitats. *Flavobacterium* was associated to all river fish and had the highest indicator value (IndVal) index (IndVal = 0.924, $P = 0.025$) meaning that it was very specific to the river and therein very abundant.

Common garden experiment

The second part of this project aimed at identifying differences in bacterial composition between different sculpin genotypes (*C. rhenanus*, invasive *C.*) when exposed to similar conditions in the laboratory. The fish were born and raised in the lab and descended from wild sculpins caught at the contact zone Sieg/Broel and brought to the lab for spawning. They were three month old when I started the experiment. Six fish of either genotype were put together in one aquarium tank, replicated by 12, so that I obtained 144 fish in total. After 3.5 month fish were killed, classified to either genotype by the amount of prickles (Koli 1969; Nolte *et al.* 2005) and the bacterial composition on their skin was investigated by high-throughput-sequencing using 16S rRNA gene markers. In the first seven weeks of the experiment many fish died due to unknown reasons, so that 53 individuals remained (21 *C. rhenanus*, 32 invasive *C.*). I sequenced the common garden samples together with those of the contact zone in two sequencing runs and identified significant variation between the runs. Further, I removed those aquarium tanks from the dataset which contained only one of both fish genotypes. Thus, only the first sequencing run remained for further statistical analyses.

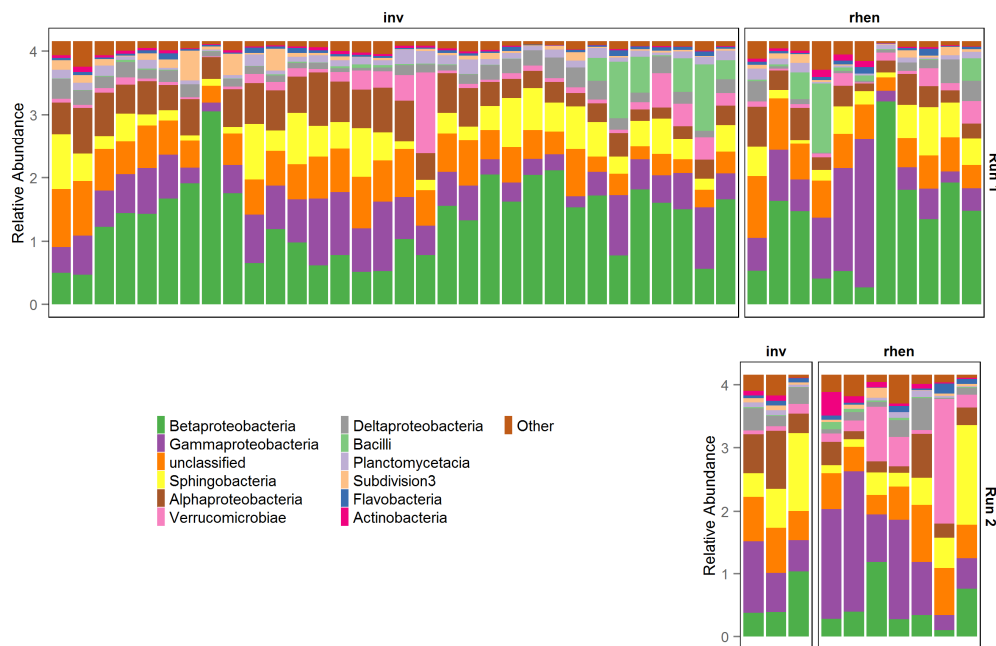


Figure 2.8: Frequency of bacterial classes of fish in common garden experiment. Shown is abundance per 2997 reads/sample. Samples are grouped according to genotype and run. Rhine (rhen) and invasive sculpins (inv) only significantly varied in the amount of Alphaproteobacteria, which was increased in invasive sculpins (KRUSKAL-WALLIS RANK SUM TEST $P = 0.025$).

Bacterial diversity in laboratory raised fish is in part shaped by host genotype

In order to determine whether the bacterial communities in Rhine and invasive sculpins vary, I first characterized the frequency of bacteria at phyla and class level and only identified significant differences between the fish genotypes in two major phyla and one major class (Figure 2.8). Overall, Proteobacteria was the most abundance phylum (inv: 62.2 %,

rhen: 58.2 %), followed by Bacteroidetes (inv: 15.1 %, rhen: 14.9 %), a proportion of unclassified reads (inv: 7.4 %, rhen: 8.4 %), Verrucomicrobia (inv: 6.6 %, rhen: 8.8 %) and Firmicutes (inv: 3.2 %, rhen: 4.6 %). At class level the bacterial frequency split into Betaproteobacteria (inv: 30.4 %, rhen: 24.0 %), Gammaproteobacteria (inv: 13.9 %, rhen: 21.7 %), a proportion of unclassified reads (inv: 13.0 %, rhen: 13.6 %), Sphingobacteria (inv: 10.7 %, rhen: 9.5 %), Alphaproteobacteria (inv: 11.6 %, rhen: 7.3 %) and Verrucomicrobiae (inv: 4.0 %, rhen: 7.0 %) In the first run, Rhine and invasive sculpins differed in the amount of Proteobacteria (KRUSKAL-WALLIS RANK SUM TEST $X^2 = 5.9383$, $P = 0.015$) and Firmicutes (KRUSKAL-WALLIS RANK SUM TEST, $X^2 = 6.3298$, $P = 0.012$). At class level the frequency distribution between the sculpin lineages significantly varied for Alphaproteobacteria, in run 1 (KRUSKAL-WALLIS RANK SUM TEST $X^2 = 5.0017$, $P = 0.025$).

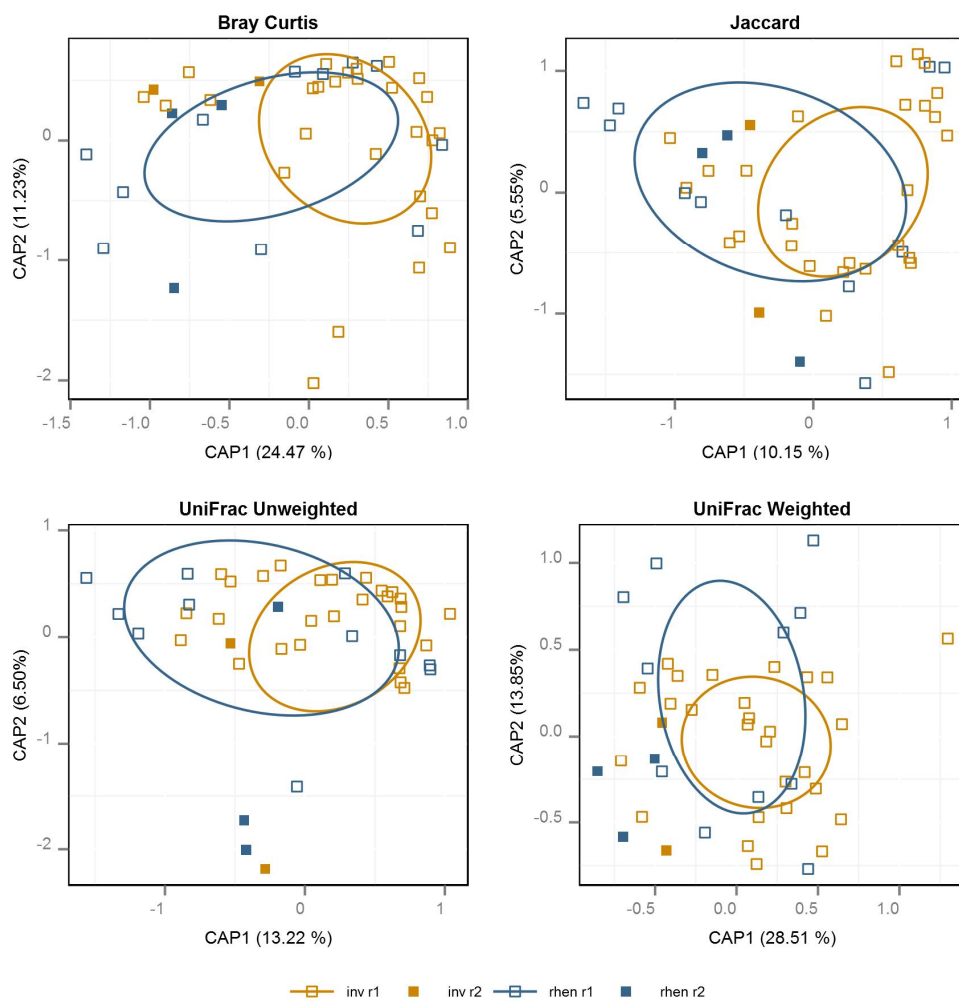


Figure 2.9: PCoA of β -diversity measurements of common garden experiment, grouped by run and genotype. Visualized are measurements of OTU abundance (Bray Curtis, weighted UniFrac) and presence/absence (Jaccard, unweighted UniFrac). Ellipses are shown for groups with $n > 3$ and denote standard deviations around the centroid of the respective group. The values were nested in tank and I obtained a significant effect for genotype in the Bray Curtis Index in the first run (ADONIS $P = 0.022$).

β -diversity analyses confirmed the results by revealing a significant influence of fish genotype (*C. rhenanus* and invasive *C.*) on Bray-Curtis dissimilarity (ADONIS $R^2 = 0.052$, $P = 0.019$) but not on any other index (Jaccard, UniFrac Weighted, Unweighted, see Table 2.S9). The ordination plot of the PCoA illustrates a separation of both fish genotypes on the first axis (Figure 2.9). This result was also supported by CCA identifying that in the best model (total inertia: 2.44858, explained inertia: 0.15165/3.83 %, conditional inertia: 1.35710/3.43 %) the significant factor genotype ($F_{1,38} = 1.7342$, $X^2 = 0.15165$, $P = 0.025$) is separating *C. rhenanus* and invasive *C.* (Table 2.S10). This difference was visualized by the graphical ordination wherein the second axis distinguishes both fish genotypes (Figure 2.10).

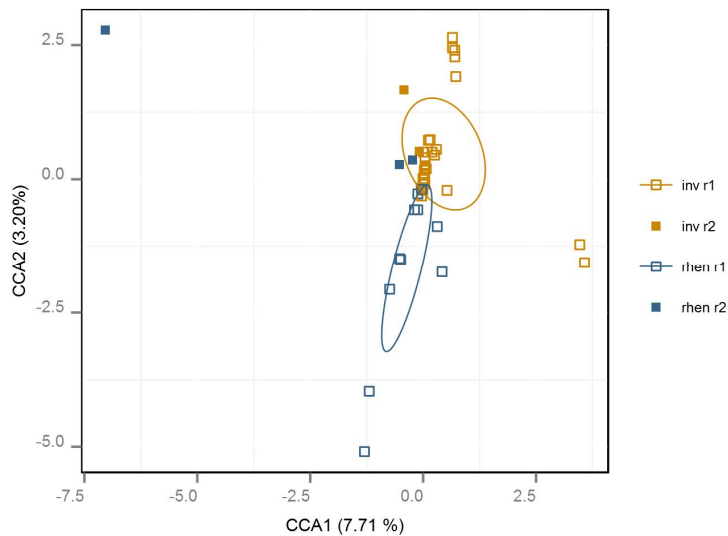


Figure 2.10: Constrained analysis of correspondence (CCA) of common garden samples, grouped by run and genotype. Axis 1 is significant (permutation test, $P = 0.023$), ellipses represent the standard deviations around the centroid of the respective genotype. CCA was conducted with tank as conditioned variable and resulted in significant effects of genotype (permutation test, $P = 0.025$).

The bacterial diversity within the fish samples (α -diversity) generated through the first sequencing run, however, determined no significant effect for fish genotype in all applied measurements (species richness, Shannon's diversity and Chao estimated richness) (Figure 2.S5; Tables 2.S7, 2.S8). Further, the indicator species analyses didn't identify any OTU significantly associated to either fish genotype ($P \geq 0.05$).

Comparison of wild caught and lab raised fish

In order to examine whether wild caught and lab raised sculpins, although they experienced very distinct environments, still have similar bacterial community structures I compared invasive and Rhine sculpins (wild) from the contact zone with those from the common garden experiment (lab). I included 35 invasive (lab_inv) and 11 Rhine sculpins (lab_rhen) from the common garden and 77 invasive (wild_inv) and 72 Rhine sculpins (wild_rhen) from the contact zone. Further, I analyzed the distinct sequencing runs separately.

Wild and lab raised sculpins harbor different bacterial communities

The phylogenetic analyses revealed in part large differences between samples of the same fish depending on the origin (wild, lab). Generally, Proteobacteria were highly abundant in lab fish vs. wild fish (43.8 % vs. 24 %, KRUSKAL-WALLIS RANK SUM TEST $X^2 = 46.969$, $P < 0.005$), whereas Verrucomicrobia were decreased (15.3 % vs. 38.4 %, KRUSKAL-WALLIS RANK SUM TEST $X^2 = 15.978$, $P < 0.005$). In the first sequencing run the influence of the origin (wild, lab) on the bacterial frequency was large: for both fish genotypes (inv, rhen) I obtained significant differences between wild and lab fish in four out of six major phyla (KRUSKAL-WALLIS RANK SUM TEST $P < 0.05$ in all cases). This was reflected at class level: again the wild and lab fish differed in four (rhen) and three (inv) major classes. In contrast, in the second run the origin had an effect on the abundance of three and two major phyla in *C. rhenanus* and invasive *C.*, respectively, and on the frequency of two and one major class(es).

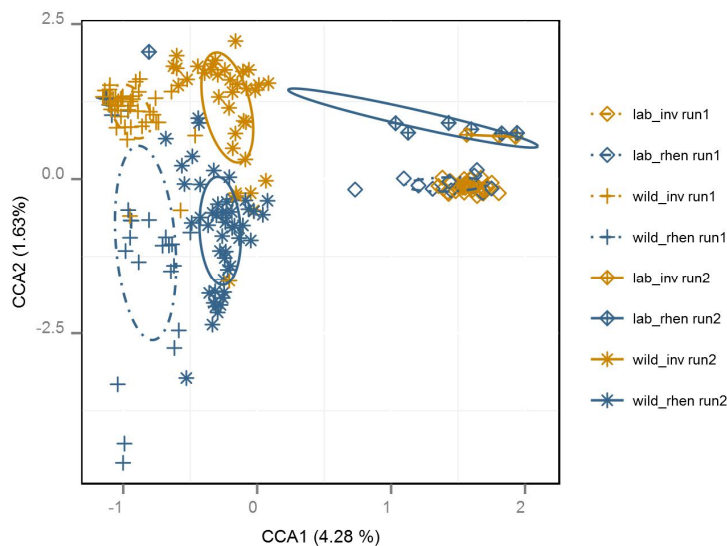


Figure 2.11: Constrained correspondence analysis with respect to origin, genotype and run. The genotypes are denoted by color (invasive *C.* = orange, *C. rhenanus* = blue), the origin (lab, wild) and runs (run1, run2) by shape. Ellipses denote standard deviation around the centroid of the mean of the respective groups. Run was included as conditioned factor. Both axes are significant (permutation test $P = 0.001$). The first axis clearly separates wild and lab fish and also the runs within the wild fish. The second axis distinguishes the fish genotypes on the wild fish and further separates the sequencing runs in the lab fish from each other. Origin and genotype were significantly influencing the variation (permutation test $P = 0.001$).

Multivariate analysis (ADONIS) of Bray-Curtis dissimilarity supported the results by revealing a significant influence of fish genotype ($P < 0.001$) and origin ($P < 0.001$) and an interaction between the factors ($P \leq 0.026$) in both runs. Also the CCA revealed a significant influence of genotype ($F_{1,196} = 2.34$, $X^2 = 0.17$, $P = 0.001$), origin ($F_{1,196} = 7.61$, $X^2 = 0.56$, $P = 0.001$) and run ($F_{1,196} = 4.17$, $X^2 = 0.31$, $P = 0.001$), as well as interactions between run and origin, run and genotype and genotype and origin (Table 2.S11). The ordination plot illustrates a clear separation of the origin on the first axis and also distinguishes wild Rhine from invasive sculpins on the second axis (Figure 2.11).

Discussion

This study aimed at investigating if bacterial microbiota is differentiated among *Cottus* genotypes and habitats, suggesting that it could act as a driver to ecological speciation processes in a fish secondary contact zone. Precisely, I here analyzed to what extent bacterial communities associated to hybridizing sculpin lineages are influenced by host genotype or environmental variation. Additionally, lab raised fish were kept in equal conditions to investigate the influence of host genotype at minimal environmental variation. At last, the comparison of wild and lab-reared fish enabled to examine if sculpins have a 'core' microbiome independent from the environment. Overall, the results suggest a substantial influence of habitat and some additional host genetic effects were also determined.

Habitat effects on skin-associated bacteria

The bacterial communities in sculpins were primarily shaped by environmental factors, which was revealed by β -diversity measurements (Table 2.S4, 2.S5) and subsequent PCoA ordination plots showing a clear separation between fish from stream and from river (Figure 2.6). Additional CCA supported the results and reflected the same pattern (Figure 2.7, Table 2.S6). Taxonomic analysis identified that fish from river contained large numbers of the phylum Verrucomicrobia which was less abundant in stream fish. Generally, the sculpin bacterial microbiota mirrored other studies on fish skin bacteria (Austin 2006; Wang *et al.* 2010; Larsen *et al.* 2013, 2015; Boutin *et al.* 2014; Chiarello *et al.* 2015; Lowrey *et al.* 2015). 2015): Proteobacteria, Firmicutes, Actinobacteria and Bacteriodetes were always the major phyla. However, only Lowrey *et al.* (2015), investigating rainbow trout, obtained the phylum Verrucomicrobia and only at low abundance. This underestimation could have been induced by primer bias, as discussed by Bergmann *et al.* (2011). The authors assessed that with the primer set 515F/806R, which was used in the present study, a good coverage can be achieved. Verrucomicrobia are known to be ubiquitous in soil and have further been isolated from freshwater, marine sediments and intertidal wetlands (Wang *et al.* 2012), as well as from grass carp eggs at fertilization stage (Wang *et al.* 2015).

When conducting studies with natural populations in the wild it is difficult if not impossible to consider all major environmental impacts affecting an individual's behavior, physiology, fitness and microbiota and this study didn't aim to examine them. Nevertheless, in the following I will discuss some considerable factors. European sculpins used to inhabit relatively cold and oxygen-rich streams (Kottelat & Freyhof 2007)) with a maximum temperature of 20°C in summer (defined as rithral) but the invasive lineage has colonized larger and oxygen-poor rivers with higher maximum temperature (>20°C) (defined as potamal). These habitats vary in several environmental factors like water chemistry, prey availability and temperature which may affect the host's microbiota. For instance, the summer temperature in the streams doesn't exceed 20 °C, whereas the temperature in the rivers can be > 20°C, as recorded by the 'Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen' (<http://luadb.lids.nrw.de/LUA/hygon/pegel.php>). Temperature is well known to influence bacterial composition in oysters (Lokmer & Wegner 2015), sponges

(Erwin *et al.* 2012) and fish (Neuman *et al.* 2016) and may affect the bacteria in sculpins, as well. Sullam *et al.* (2012) revealed that bacterial composition in fish also depends on salinity and host trophic level but I didn't expect these factors to cause variation in sculpin-associated bacteria. The salinity is most likely very similar in both habitats and even if the overall composition of invertebrates, the major food of sculpins, differs in rivers and streams, the general trophic level remains.

Overall, these findings agree with other studies, for instance identifying that microbial communities in mice are shaped by geography (Linnenbrink *et al.* 2013) and skin bacteria of amphibians are site-specific (Kueneman *et al.* 2014). It has been hypothesized amongst others by Boutin *et al.* (2013) and Fitzpatrick & Allison (2014) that the host's skin is possibly colonized by bacteria from the surrounding environment. Contrary to that are findings documenting that hosts and environmental bacteria are distinct and therefor discuss that it's rather the host which is driving the bacterial composition (Wong & Rawls 2012; Walke *et al.* 2014; Larsen *et al.* 2015). To uncouple environmental from host genetic factors I performed a common garden experiment in which I controlled for the environment so that any effects could be related to genetic differences. Results are discussed further down.

Genetic effects on skin associated bacteria

Apart from habitat also the host genotype had a significant effect on bacterial community structure, albeit the impact was minor compared to habitat effects. The taxonomic analysis revealed significant differences between all three fish genotypes in the stream; however, the hybrids accounted for most of the variation (see phylogenetic analysis pp. 10). Recent studies on microbiota in hybrids revealed that compared to the parents the hybrids' bacterial composition is irregular (Brucker & Bordenstein 2013) and divergent (Wang *et al.* 2015). Generally, hybrids can have extreme or novel phenotypes, exceeding those of their parents, which is referred to as transgressive segregation (Rieseberg *et al.* 1999a, 2003b; Nolte & Sheets 2005). This phenomenon occurs due to new genetic combinations generated by recombination. A well described example for transgressive segregation is the emergence of three hybrid sunflower species which resulted from hybridization of *Helianthus annuus* and *H. petiolaris* (Rieseberg *et al.* 1996, 2003a). Additive and epistatic gene combinations make the hybrid species superior to their parents and enable them to occur at extreme habitats. Transgressive segregation in the sculpin hybrids have been determined before (Nolte *et al.* 2005; Czypionka *et al.* 2012) but the pattern in the current study is less pronounced and has to be further investigated for instance by quantitative trait loci (QTL) analysis. Apart from phylogenetic analysis also β -diversity metrics determined a significant influence of host genotype on bacterial composition (ADONIS, $P = 0.001$) (Table 2.S4, 2.S5), as well as CCA (permutation test $P = 0.033$) (Table 2.S6). However, all three sculpin lineages were very dispersed in the ordination plots and no clear clustering was apparent (Figure 2.6, Figure 2.7).

Host genetic effects on bacterial microbiota have been assessed in a variety of studies for instance in humans (e.g. Turnbaugh *et al.* 2009; Goodrich *et al.* 2014), mice (e.g. Benson *et*

al. 2010; Spor *et al.* 2011; Campbell *et al.* 2012) and fish (e.g. Rawls *et al.* 2006; Roeselers *et al.* 2011; Boutin *et al.* 2014). There's evidence for a core intestinal microbiota independent of the environment (Roeselers *et al.* 2011). Subsequently, related individuals share more bacteria than non-related ones (Turnbaugh *et al.* 2009). However, family members often share similar lifestyles, diets and environments and thus genetic effects are not independent (Song *et al.* 2013). A direct effect on microbiota is exerted by several host genes, most of them are parts of the immune system and involved in metabolism (reviewed by Spor *et al.* 2011). Toll-like-receptors (TLR) play a key role in the innate immune system, they detect molecules derived from microbes. A study revealed that lacking TLR5 altered caecal microbiota in mice and caused metabolic syndrome (Vijay-Kumar *et al.* 2010).

One aim of this study was to assess whether bacterial microbiota in sculpins are considerably influenced by host genotype and can thus induce reproductive barriers between individuals and drive adaptive speciation. There are several studies determining that microbes can cause reproductive isolation (reviewed by Brucker & Bordenstein 2012). Some pea aphids (*Acyrtosiphon pisum*) have a bacterial symbiont which induces a fitness advantage on white clover plants (*Trifolium repens*) compared to their conspecifics lacking the symbiont (Tsuchida *et al.* 2004). This caused niche extension and geographic isolation. Nolte (Nolte *et al.* 2006) identified that in the sculpins contact zone selection acts against immigrants and that there are no intrinsic incompatibilities. In one possible scenario the parental lineages, Rhine and invasive sculpins are both adapted to local bacteria, be they pathogens, mutualists or commensals, acting as selective force and preventing dispersal to new habitats. In the hybrids new combination of genes can lead to reduced fitness in both habitats stream and river and thus prevents the emergence of a new hybrid lineage (Burke & Arnold 2001).

Common Garden Experiment

In the common garden experiment *C. rhenanus* and invasive *C.* were raised and kept under the same conditions so that environmental variation was minimized and any resulting variation was related to host genetic effect. Like in the contact zone, I identified differences between the genotypes in the lab fish, as well, supporting the idea that the host genotype in part shapes the bacterial composition. The invasive sculpins had increased species richness and significantly differed from the Rhine sculpins in Bray Curtis β -diversity index ($R^2 = 0.051$, $P = 0.019$) and the CCA supported the results (Table 2.S10). In the beginning of the experiment in all tanks fish died for unknown reason so I lost many replicates. Hence, the statistical power in all tests was reduced and the obtained results could be biased by an unknown factor. If the cause of death was an infection, the fish skin could contain an increased amount of pathogenic bacteria which may have caused the disease. Contrary, the skin could be colonized by commensals protecting mucosal epidermis against pathogens (reviewed by Lazado & Caipang 2014). The taxonomic analyses didn't reveal a remarkable increase in known fish pathogens or commensals but since there's a large amount of unknown and unclassified OTUs, I cannot certainly conclude there are none. However, after 2/3 of the duration of the experiment (108 days in total) no fish died any-

more and at sampling time they all looked healthy and I assume no considerable bias. In sum, even though the sample size was small and potentially influenced by an unknown factor the common garden experiment illustrates that in lab raised sculpins the host genotype has an influence on the bacterial community which is independent from the environment.

Comparison of wild vs. lab fish

In order to investigate whether fish from the wild and from the lab, albeit exposed to very distinct environments and food, still have similar bacterial communities, I compared invasive and Rhine sculpins from the contact zone with those from the common garden experiment. Interestingly, I obtained a significant influence of genotype in Bray Curtis dissimilarity analysis in both runs (ADONIS, $P < 0.001$), as well as in the constrained correspondence analysis (Table 2.S11). This result was clearly visible in the ordination plot, either (Figure 2.11) and suggests that sculpins have a core skin microbiota which is shaped by selective pressures within the fish habitat being independent from the environment. Similar results have been demonstrated for zebrafish (*Danio rerio*): lab-reared fish have a gut bacterial community similar to that of their wild relatives (Roeselers *et al.* 2011). Other studies analyzed differences between wild and lab hydra (Fraune & Bosch 2007), fruitflies (Staubach *et al.* 2013) and mice (Wang *et al.* 2015) and all identified that lab-reared individuals, sometimes kept in controlled facilities for many generations, have microbial compositions which reflect that of their wild conspecifics. A possible explanation for this similarity can be that the sculpins still have the same physiology, anatomy and immune systems, if they grew up in the wild or in the lab.

Additional to host genotype, most of the differences in any analysis could be related to origin (wild, lab). Lab and wild fish shared only 16 % of all OTU and differed in all major taxa and orders. Further, origin explained 17.5 % of the variation in the Bray-Curtis dissimilarity analysis and subsequent CCA revealed a significant influence, either, which was also apparent on the ordination plot (Figure 2.11). In nature, organisms are exposed to numerous biotic and abiotic factors and all shape their bacterial microbiota, as discussed earlier. In contrast, laboratory environments are artificial and the influencing factors are restricted, thus, variation between lab and wild sculpins were expected. Additionally, DNA of wild and lab-reared sculpin fins was extracted with different protocols and thereafter stored for 8 years (wild samples) or only several weeks (lab samples), respectively, before sequencing. Several studies have shown that the DNA extraction is greatly influencing the outcome of a bacterial study (Yuan *et al.* 2012; Kennedy *et al.* 2014; Rubin *et al.* 2014; Larsen *et al.* 2014b). Hence, this might have had an effect on my results, either.

Sequence run variation

In the current study, samples have been sequenced in two different libraries with different kits, each, which caused significant variation in bacterial communities. Thus, I included the factor run (run1, run2) as main effect in linear models or as condition in partial constrained analysis. Further, I applied statistical tests for each run, separately, when examining the

bacterial frequency distribution or β -diversity indices. It is evident, that sequence data quality, read abundance and classification depend for instance on the primer pairs (Mao *et al.* 2012), DNA extraction protocol (Yuan *et al.* 2012; Kennedy *et al.* 2014; Rubin *et al.* 2014; Larsen *et al.* 2014b), library preparation (Jones *et al.* 2015), the sequencing facility and the sequencer itself (Kim & Yu 2014) as well as on the analysis pipeline and taxonomic database (Mizrahi-Man *et al.* 2013). To directly compare the libraries with each other I included 15 technical replicates, DNA samples which were sequenced in both sequencing runs. I didn't obtain significant effects of run in any analysis. Since the methods for both library preparations were generally the same, I hypothesize that the variation in our data can possibly be explained by the use of different sequencing kits. The 15 technical replicates, which were included in both libraries, did not statistically differ in any analyses but since the whole dataset contained 16 times more samples, it has more statistical power. Results including samples from different libraries or studies should be interpreted with care, which has recently been confirmed by Lozupone *et al.* (2013). The authors conducted meta-analyses of human microbiota from several studies all generated with the same 16S rRNA gene sequences. They discovered that fecal samples from Western adults clustered by study and not by body site or age. It is necessary and important to consider all parameters that may influence the microbiome and to carefully standardize all experimental protocols to make the results reproducible (reviewed by Laukens *et al.* 2016). Since more and more protocols, Kits and sequencing platforms are developed, it is particularly important to understand the possible type of bias in the data and to optimize techniques to minimize these biases (Lozupone *et al.* 2013).

In sum, although sequencing run explained a large amount of the data variation I generally identified the same or similar patterns within each run as well as in the whole dataset. Thus, I assume that the results would be clearer without the run discrepancy but would generally not change.

Conclusion

This study about skin microbiota in different fish lineages, inhabiting a naturally emerged secondary contact zone, provides important insight to understand the role of environmental and genetic factors in shaping bacterial communities. I revealed a major influence of host habitat and an additional impact of host genetic architecture on bacterial composition of sculpin skin. Similarities between wild and lab-reared fish further indicate that there's selection pressure acting on the fish skin to select for certain bacteria. Considering the results, I'm not able to conclude that bacterial microbiota plays an essential role in driving adaptive speciation of sculpins because migrating fish could quickly adapt to the habitat-specific bacterial communities. Nevertheless, since no intrinsic reproductive barrier has been determined, the environment surely plays a main role in the emergence of reproductive isolation and acts as selective force on the hybrid fish and their parents. In theory, the hybrid zone enables to separate environmental from genetic effects since the same fish genotype occurs in two distinct habitats and different fish genotypes are found in the same habitat. However, the contact zone is narrow and dispersal to foreign habitats is limited.

Only few pure Rhine sculpins were found in the stream and few pure invasive sculpins were found in the river, hence this system lacks statistical power. Thus, a key result of this study is that the genetic structure in the contact zone didn't permit conclusive tests to answer the overall question: whether host-associated bacteria act as a driver to adaptive speciation processes. Future work would need to study fish species occurring in both habitats stream and river. Thereby, it would be possible to determine whether the strong habitat effect I found for the sculpins is also visible in other species.

Supplement

Figures

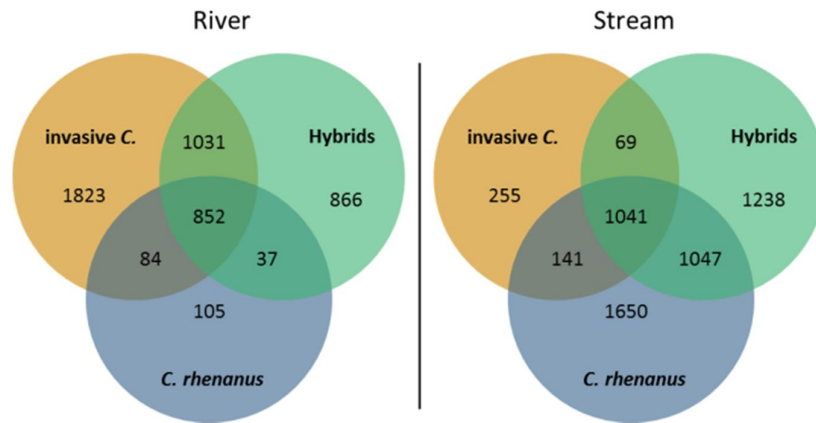


Figure 2.S1: Venn diagram illustrating the number of shared OTUs in the respective genotypes per habitat stream or river.

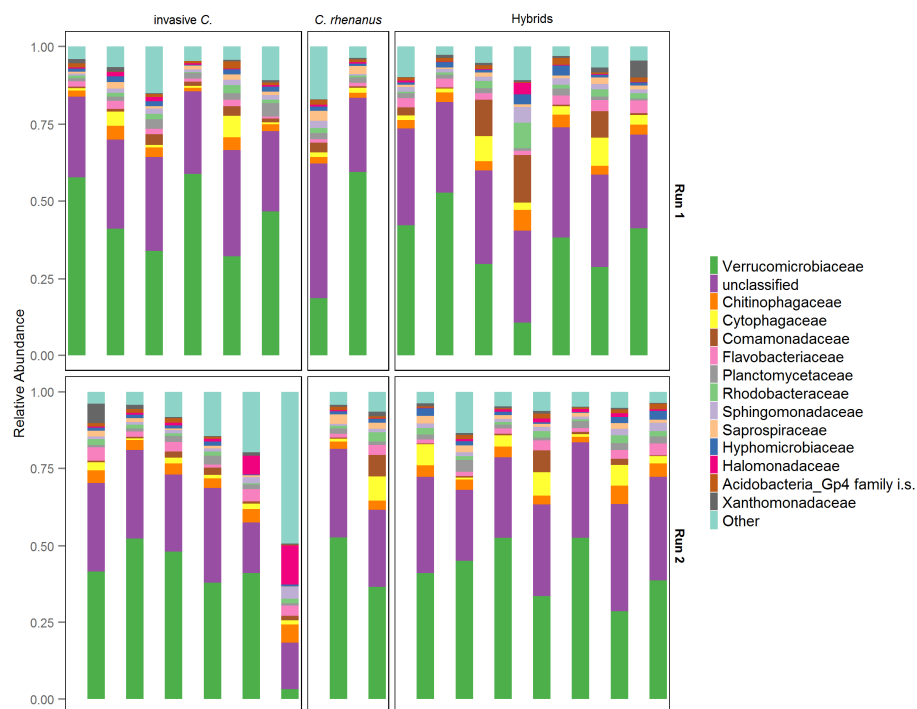


Figure 2.S2: Relative read abundance of 15 technical replicates at phylum level. I included 15 samples of all three genotypes in both sequencing runs and phylogenetic analyses didn't reveal a significant effect of run in any phylum (KRUSKAL-WALLIS RANK SUM TEST, $P \geq 0.05$).

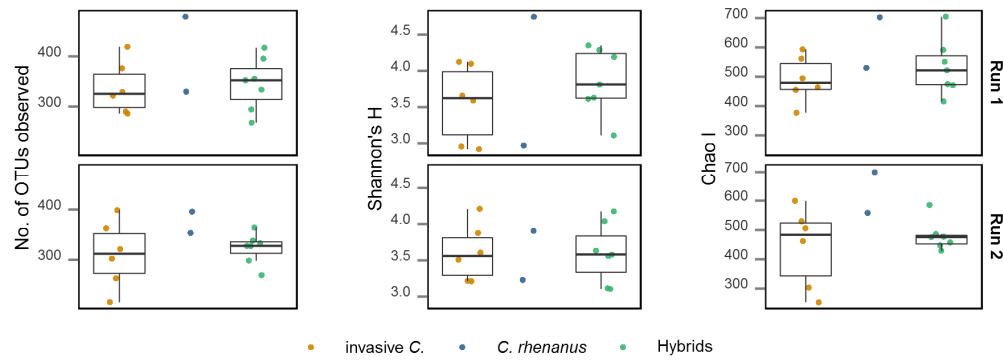


Figure 2.S3: α -diversity metrics of technical replicates. I calculated species richness, diversity and estimated richness and linear models revealed no significant effects of sequencing run ($P \geq 0.05$) in any index.

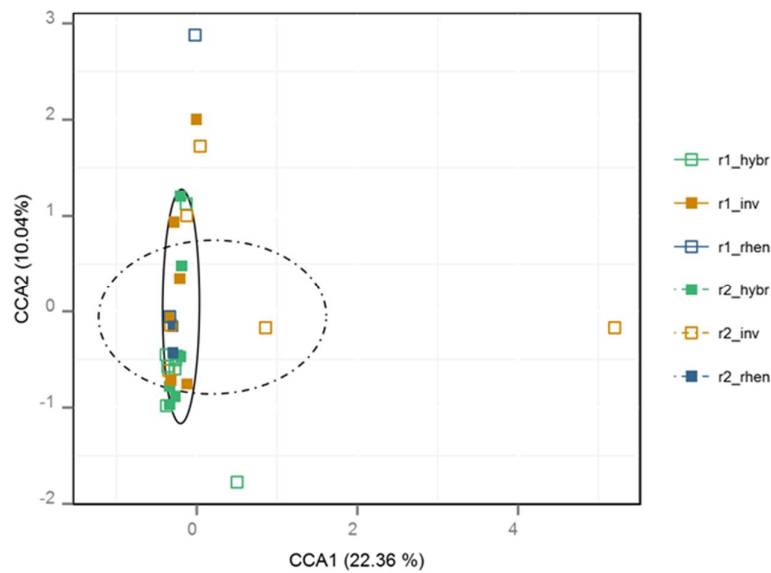


Figure 2.S4: Constrained correspondence analysis of technical replicates. Species are grouped by genotype (rhen, hybr, inv) and run (r1, r2). Ellipses denote standard deviations around the centroid of the respective group. ANOVA-like permutations identified no significant effect of any factor ($P \geq 0.05$).

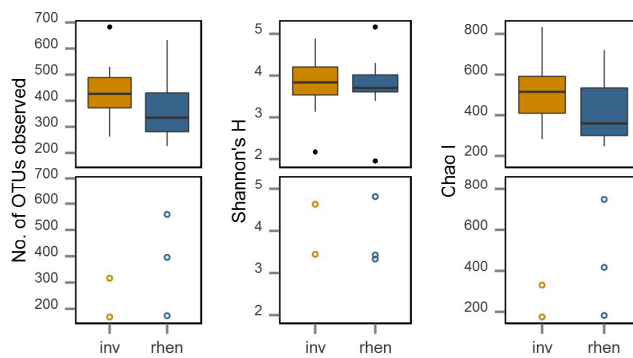


Figure 2.S5: α -diversity of common garden experiment. Illustrated are bacterial richness (no. of observed OTUs), diversity (Shannon's H) and estimated richness (Chao I) of invasive *C.* (*inv*) and *C. rhenanus* (*rhen*). Boxplots are only shown for groups with $n \geq 5$. Statistic was only applied on samples in run1 and no significant differences between the genotypes were identified.

Tables

Table 2.S1: Analysis of CCA of technical replicates. Included terms are genotype (*inv*, *hybr*, *rhen*) and run (*run1*, *run2*). Significance of model, terms and axes was evaluated by using ANOVA-like permutation tests.

	Df	ChiSquare	F	Pr(>F)
Significance of Model				
Model	5	0.39525	1.0285	0.308
Residual	24	1.84472		
Significance of Model terms				
Genotype	2	0.16533	1.0755	0.274
Run	1	0.06992	0.9097	0.696
genotype:run	2	0.16000	1.0408	0.323
Residual	24	1.84472		
Significance of axes				
CCA1	1	0.15568	2.0254	0.029
CCA2	1	0.09045	1.1767	0.243
CCA3	1	0.06823	0.8877	0.620
CCA4	1	0.04240	0.5517	0.981
CCA5	1	0.03849	0.5008	0.990
Residuals	24	1.84472		

Bold P-values indicate significance ($P < 0.05$).

Table 2.S2: Linear model of α -diversity in the contact zone with significance of model terms. The full model contained the terms: fish genotype (rhen, inv, hybr), sequencing run (run1, run2) and habitat (stream, river); shown is the best model selected by minimizing the Akaike Information Criterion (AIC) value at lowest model complexity.

Index	Term	Df	Sum Sq	Mean Sq	F value	Pr (>F)
(Species richness) ²	Run	1	1892.0	1891.97	123.02	<2e-16
	Residuals	246	3783.4	15.38		
sqrt(Chao)	Run	1	4430.2	4430.2	198.14	< 2.2e-16
	Residuals	250	5589.8	22.4		
(Shannon) ²	Run	1	686.1	686.09	15.6225	0.000101
	Habitat	1	6	5.98	0.1361	0.712474
	run:habitat	1	1439.8	1439.85	32.786	2.96e-08
	Residuals	248	10891.3	43.92		

Bold P-values indicate significance ($P < 0.05$).

Table 2.S3: Linear model of α -diversity indices in the contact zone with significance of coefficients. The Intercept indicates the habitat river in run1 for species richness and Shannon and run1 for Chao.

Index	Term	Estimate	Std. Error	t value	Pr(> t)	adj. R ²
(Species richness) ²	(Intercept)	19.4154	0.3774	51.45	<2e-16	0.3307
	Run (Run2)	-5.5707	0.5023	-11.09	<2e-16	
sqrt(Chao)	(Intercept)	23.6066	0.4509	52.36	<2e-16	0.4399
	Run (Run2)	-8.4542	0.6006	-14.08	<2e-16	
(Shannon) ²	(Intercept)	14.569	0.781	18.655	< 2e-16	0.1536
	Run (Run2)	1.643	1.259	1.305	0.193	
	Habitat (Stream)	5.984	1.329	4.504	1.03e-05	
	Run (Run2):Habitat (Stream)	-10.234	1.787	-5.726	2.96e-08	

Bold P-values indicate significance ($P < 0.05$).

Table 2.S4: Permutation multivariate analysis of dissimilarity matrices in the contact zone. Terms are explained in Table 2.S2, interactions between terms are denoted with a colon.

Index	Terms	Df	SumsSqs	MeanSqs	F. Model	R ²	Pr (>F)
Bray Curtis	Genotype	2	3.461	1.7304	8.435	0.04861	0.001
	Habitat	1	8.027	8.0267	39.126	0.11273	0.001
	Run	1	7.19	7.1902	35.048	0.10098	0.001
	Genotype:Habitat	2	0.57	0.2849	1.389	0.008	0.125
	Genotype:Run	2	1.03	0.515	2.51	0.01446	0.006
	Habitat:Run	1	1.348	1.3481	6.571	0.01893	0.001
	Genotype:Habitat:Run	2	0.341	0.1704	0.83	0.00479	0.667
	Residuals	240	49.236	0.2052		0.6915	
Total	251	71.203			1		
Jaccard	Genotype	2	1.688	0.844	2.6033	0.01898	0.001
	Habitat	1	3.162	3.1621	9.7531	0.03555	0.001
	Run	1	3.563	3.5635	10.9912	0.04006	0.001
	Genotype:Habitat	2	0.693	0.3464	1.0684	0.00779	0.233
	Genotype:Run	2	0.866	0.4331	1.3359	0.00974	0.02
	Habitat:Run	1	0.619	0.6187	1.9082	0.00695	0.006
	Genotype:Habitat:Run	2	0.553	0.2767	0.8536	0.00622	0.93
	Residuals	240	77.811	0.3242		0.87472	
Total	251	88.956			1		
Unifrac unweighted	Genotype	2	0.01821	0.009103	8.313	0.03678	0.001
	Habitat	1	0.05763	0.057632	52.63	0.11643	0.001
	Run	1	0.14397	0.143969	131.474	0.29084	0.001
	Genotype:Habitat	2	0.00087	0.000435	0.398	0.00176	0.792
	Genotype:Run	2	0.00782	0.00391	3.571	0.0158	0.023
	Habitat:Run	1	0.00286	0.002859	2.611	0.00578	0.084
	Genotype:Habitat:Run	2	0.00085	0.000423	0.386	0.00171	0.779
	Residuals	240	0.26281	0.001095		0.53091	
Total	251	0.49501			1		
Unifrac weighted	Genotype	2	0.2575	0.12874	16.314	0.07071	0.001
	Habitat	1	0.7428	0.7428	94.131	0.20399	0.001
	Run	1	0.5445	0.54449	69.001	0.14953	0.001
	Genotype:Habitat	2	0.0194	0.00969	1.228	0.00532	0.283
	Genotype:Run	2	0.0752	0.0376	4.765	0.02065	0.002
	Habitat:Run	1	0.103	0.10301	13.054	0.02829	0.001
	Genotype:Habitat:Run	2	0.0052	0.00258	0.327	0.00142	0.902
	Residuals	240	1.8939	0.00789		0.52009	
Total	251	3.6414			1		

Bold P-values indicate significance ($P < 0.05$).

Table 2.S5: Permutation multivariate analysis of dissimilarity matrices in the contact zone per run. Terms are explained in Table 2.S2.

Index	Terms	Df	SumsSqs	MeanSqs	F. Model	R ²	Pr (>F)	
Bray Curtis	genotype	2	1.6403	0.8202	5.0427	0.07266	0.001	
	habitat	1	3.6944	3.6944	22.715	0.16365	0.001	
	run 1	Residuals	106	17.2401	0.1626		0.76369	
		Total	109	22.5748			1	
	run 2	genotype	2	1.655	0.82736	3.474	0.04478	0.001
		habitat	1	2.43	2.43016	10.204	0.06577	0.001
		Residuals	138	32.866	0.23816		0.88945	
		Total	141	36.951			1	
Jaccard	genotype	2	1.0336	0.51681	1.9099	0.03293	0.001	
	habitat	1	1.6723	1.67231	6.1802	0.05328	0.001	
	run 1	Residuals	106	28.6827	0.27059		0.91379	
		Total	109	31.3887			1	
	run 2	genotype	2	1.186	0.59323	1.626	0.02251	0.001
		habitat	1	1.183	1.18323	3.2431	0.02244	0.001
		Residuals	138	50.349	0.36484		0.95505	
		Total	141	52.718			1	
Unifrac unweighted	genotype	2	0.9028	0.45138	2.0664	0.03547	0.001	
	habitat	1	1.3912	1.39116	6.3688	0.05467	0.001	
	run 1	Residuals	106	23.1541	0.21843		0.90986	
		Total	109	25.448			1	
	run 2	genotype	2	0.858	0.42908	1.599	0.02217	0.005
		habitat	1	0.822	0.82222	3.0641	0.02124	0.001
		Residuals	138	37.031	0.26834		0.95659	
		Total	141	38.711			1	
Unifrac weighted	genotype	2	0.3858	0.19291	6.4318	0.08578	0.001	
	habitat	1	0.9328	0.93281	31.1001	0.20738	0.001	
	run 1	Residuals	106	3.1794	0.02999		0.70684	
		Total	109	4.498			1	
	run 2	genotype	2	-0.04176	-0.02088	-2.9249	-0.04269	0.949
		habitat	1	0.0348	0.034797	4.8749	0.03558	0.062
		Residuals	138	0.98503	0.007138		1.00711	
		Total	141	0.97807			1	

Bold P-values indicate significance ($P < 0.05$).

Table 2.S6: Analysis of partial CCA of OTUs from contact zone samples. The included terms are habitat and genotype with run as conditioned variable. Significance of model, terms and axes was evaluated by using permutation tests. Both factors had a significant effect on the data variation.

	Df	ChiSquare	F	Pr(>F)
Significance of Model				
Model	3	0.4098	2.3348	0.001
Residual	247	14.4508		
Significance of Model terms				
Habitat	1	0.2761	4.7201	0.001
Genotype	2	0.1336	1.1422	0.032
Residual	247	14.4508		
Significance of axes				
CCA1	1	0.2762	4.7207	0.001
CCA2	1	0.0737	1.2591	0.053
CCA3	1	0.0599	1.0246	0.440
Residuals	247	14.4508		

Bold P-values indicate significance ($P < 0.05$).

Table 2.S7: Linear model of α diversity in the common garden experiment with significance of model terms. I calculated the mean value per tank and kept only those containing both genotypes, so I applied linear models on 7 replicates, each. Run2 was excluded from the data set because therein the groups contained not enough samples ($n \leq 3$). The full model contained the term fish genotype (rhen, inv); shown is the best model selected by minimizing the Akaike's Information Criterion (AIC) value at lowest model complexity.

Index	Term	Df	Sum Sq	Mean Sq	F value	Pr (>F)
log(Species richness)	Residuals	14	0.9916	0.070829		
sqrt(Chao)	Residuals	14	1.2991	0.092789		
log(Shannon)	Residuals	14	0.36956	0.026397		

Bold P-values indicate significance ($P < 0.05$).

Table 2.S8: Linear model of α -diversity in common garden experiment with significance of coefficients. For detailed description see Table 2. S7.

Index	Term	Estimate	Std. Error	t value	Pr(> t)	adj. R ²
Species richness	(Intercept)	6.03662	0.06872	87.85	<2e-16	
sqrt(Chao)	(Intercept)	6.19775	0.07865	78.8	<2e-16	
log(Shannon)	(Intercept)	1.55929	0.04195	37.17	2.15e-15	

Bold P-values indicate significance ($P < 0.05$).

Table 2.S9: Permutation multivariate analysis of dissimilarity matrices in the common garden experiment. The included term was genotype (rhen, inv). Permutations were applied within each tank.

Index	Terms	Df	SumsSqs	MeanSqs	F. Model	R ²	Pr (>F)
Bray Curtis	Genotype	1	0.3527	0.35272	1.9656	0.05177	0.022
	Residuals	36	6.4601	0.17945		0.94823	
	Total	37	6.8128			1.00000	
Jaccard	Genotype	1	0.4005	0.40054	1.4963	0.03991	0.077
	Residuals	36	9.6368	0.26769		0.96009	
	Total	37	10.0374			1.00000	
Unifrac unweighted	Genotype	1	0.0672	0.067203	1.3513	0.03618	0.25
	Residuals	36	1.7903	0.049731		0.96382	
	Total	37	1.8575			1.00000	
Unifrac weighted	Genotype	1	0.2397	0.23971	1.2274	0.03297	0.097
	Residuals	36	7.0309	0.19530		0.96703	
	Total	37	7.2706			1.00000	

Bold P-values indicate significance ($P < 0.05$).

Table 2.S10: Canonical correspondence analysis of OTUs of common garden experiment samples. The included terms are explained in Table 2.S7. Significance of model, terms and axes was evaluated by using permutation tests.

	Df	ChiSquare	F	Pr(>F)
Significance of Model				
Model	1	0.15165	1.7342	0.019
Residual	28	2.44858		
Significance of Model terms				
genotype	1	0.15165	1.7342	0.025
Residual	28	2.92569		
Significance of axes				
CCA1	1	0.15165	1.7342	0.025
Residuals	28	2.92569		

Bold P-values indicate significance ($P < 0.05$).

Table 2.S11: Canonical correspondence analysis of OTU of wild vs. lab sculpins. The included terms are origin (wild, lab) and genotype (inv, rhen) and permutations were applied within the runs. Significance of model, terms and axes was evaluated by using permutation tests.

	Df	ChiSquare	F	Pr(>F)
Significance of Model				
Model	3	0.9699	4.7857	0.001
Residual	198	13.3754		
Significance of Model terms				
origin	1	0.6110	9.0449	0.001
genotype	1	0.1990	2.9459	0.001
Residual	198	0.1598	2.3662	0.001
Significance of axes				
CCA1	1	0.6142	9.0919	0.001
CCA2	1	0.2343	3.4681	0.001
CCA3	1	0.1214	1.7971	0.058
Residual	198	13.3754		

Bold P-values indicate significance ($P < 0.05$).

CHAPTER 2

Comparative analysis of skin-associated bacteria in teleost fish inhabiting a secondary contact zone of *Cottus*

Introduction

Relationships between eukaryotes and bacteria are probably as old as the eukaryotes themselves (McFall-Ngai *et al.* 2013; Alegado & King 2014). Over time the role of bacteria changed from being 'just prey' to becoming essential mutualists for their host with specialized functions and it becomes clear that they are more than just 'transient passengers' acquired from the environment or inherited from mother to offspring by vertical transfer. They provide digestible nutrients in the intestinal tract of their hosts (McFall-Ngai *et al.* 2013), influence their host's development (Fraune & Bosch 2010) and behavior (Cryan & Dinan 2012) and stimulate its immune system (Sekirov *et al.* 2010). A healthy bacterial microbiota is very specific and relatively stable over time (Costello *et al.* 2009) and many taxa are endemic to their hosts. Gut bacteria in termites, for instance, represent monophyletic clusters being distinct from other bacterial lineages and their composition greatly differs between termite host genera (Schmitt-Wagner *et al.* 2003; Hongoh *et al.* 2005), suggesting a longtime co-evolution. The outcome of symbiotic co-evolution is influenced e.g. by host diet (Ley *et al.* 2008), physiology (Hooper *et al.* 2012) and lifestyle (Song *et al.* 2013) but also by environmental factors, like salinity (Lozupone & Knight 2007), temperature (Lokmer & Wegner 2015) and seasonality (Crump & Hobbie 2005). The entity of a host and its microbial symbionts was summarized as holobiont (Margulis 1993), a unit selections acts on, and hosts were no longer seen as just being environments for their microbes but as genotypes interacting with their associated bacterial genomes and with the environment surrounding them (Bordenstein & Theis 2015). These tight and specialized relationships between host and microbes can lead to ecological isolation between locally adapted host populations having and lacking certain bacteria. Hybridization can then cause a breakdown of genomic structures in the holobiont and thus reproductive isolation can arise (Brucker & Bordenstein 2012).

The significance of bacteria to their hosts is beyond doubt, however, the importance of microbes in host speciation processes has been rediscovered only recently (Brucker & Bordenstein 2013; Wang *et al.* 2015). Both studies demonstrated a reduced fitness or even inviability of hybrids between two species of *Nasonia* wasps and house mice, respectively. However, there's need to study more such systems to gain better insights into processes underlying speciation by symbiosis. Despite using artificial hybrids in controlled laboratory conditions it is acutely evident to investigate the emergence of bacteria-induced reproductive barriers under positive or negative selection in the wild. However, the holobiont consists of three components: the host, the bacteria and the environment, which makes it difficult to disentangle the contribution to diverging evolution of the particular factors. In

principle, naturally emerged hybrid zones should enable to separate the components of this triangular relationship since they often emerge at ecotones and thus distinct genotypes, the hybridizing parents and their admixed offspring, occur in different habitats. One can thus compare distinct host genotypes in the same environment and detect differences of the same host genotype in different environments. The system permits to test one factor while controlling for the other one. Hybrid zones have been extensively used for genetic mapping because therein thousands of generations of recombinants are generated under natural selection and enables to study the genetic basis of adaptation in non-model organisms which are usually not amenable to experimental crosses (Buerkle & Lexer 2008).

Different lineages of European freshwater sculpins (*Cottus* sp.) are hybridizing in several tributaries of the Rhine system near Bonn, Germany, and thus provide an excellent system to study early evolutionary processes and to determine the impact of extrinsic factors, like bacteria or environmental factors, on adaptive speciation. Presumably 200 years ago man-made canals connected the river systems Rhine and Scheldt which caused hybridization between the sculpins *Cottus rhenanus* and *C. perifretum* from small tributaries. The resulting hybrids segregated from their parents and invaded new habitats which have been free of sculpins until then. Hence, they are in the following referred to as invasive *Cottus*. In the 1980s, the invasive sculpins have constantly been found in the less oxygenated and warmer river Rhine and its major tributaries (Nolte *et al.* 2005). Meanwhile, they exceeded their range which led to the emergence of several, narrow secondary contact zones with one of their parents, the Rhine sculpins, about 20 years ago (Nolte *et al.* 2006). The contact zones are found where small tributaries disemboque into larger rivers and therein, both lineages hybridize under absence of intrinsic barriers but apparently environmental selection acts upon immigrant genotypes, they seem to be restricted to their primordial location (Nolte *et al.* 2006, 2009). Additionally, the hybrid dispersal is limited and selection prevents their adaptation to new environments, either.

In a previous study I investigated bacterial communities on sculpin fins sampled at 10 sites along the contact zone Sieg/Broel near Bonn, Germany (Figure 2.1). Pure parental fish as well as admixed hybrids are migrating into foreign habitats but the hybrid zone is only 2 km width and dispersal is limited (Nolte *et al.* 2006). The fish were genotyped for 85 SNP markers (Stemshorn 2007) to assess their genetic background and hereupon classified into three groups: *Cottus rhenanus*, hybrids and invasive *Cottus* (see Chapter I). The contact zone study was completed by a common garden experiment wherein *C. rhenanus* and invasive *C.* were raised and kept in the lab under equal conditions to examine the host genetic impact on the bacterial composition under controlled and akin environment. The results revealed a major influence of habitat on the fish microbiota which is in line with previous studies (Spor *et al.* 2011; Sullam *et al.* 2012; Linnenbrink *et al.* 2013). Additionally, I identified an influence of host genotype which reflects the results of other publications (Larsen *et al.* 2013; Walke *et al.* 2014), as well. Subsequently, I detected that in wild caught as well as in lab raised sculpins that microbiota associated with host genotypes, that is invasive *C.* and *C. rhenanus*, clearly differ. That gives raise to the assumption that some separating factors are independent from the environment. In the contact zone several pure

Rhine sculpins are found in the larger river Sieg and at the same time invasive sculpins are found in the small tributary Broel. However, the analysis of the genetic structure of this hybrid zone (chapter I) revealed that migration and dispersal is limited by natural selection. As a consequence, given the sample numbers available here, the fish genotypes are not sufficiently independent from the environment. In other words, I have found too few immigrant genotypes in the foreign habitats to perform the extensive statistical analysis that this system permits in principle. To still address some of the general questions of this thesis an extension to additional fish species, occurring in both habitats and do not genetically vary, is needed.

This enables to detach host genetic from environmental effects and would determine if habitat effects, like in sculpin-associated microbes, are a general phenomenon shared by other fish species. Therefore, I investigated the bacterial microbiota of the stone loach (*Barbatula barbatula*) and the common minnow (*Phoxinus phoxinus*), both distributed in rivers and streams throughout Europe (Kottelat & Freyhof 2007) and also inhabit stream and river of the contact zone Sieg/Broel. I further assume that stream and river populations have the same genetic architecture as there is currently no evidence that suggests otherwise. In order to relate the minnows and stone loaches skin microbiota to that of *Cottus* spp. and further to compare bacterial communities of previously caught sculpins with that of recently collected ones we caught invasive sculpins in the Sieg and Rhine sculpins in the Broel.

In sum, the objectives of this study were: (i) to assess the bacterial microbiota of the distinct fish lineages stone loach, minnow and sculpin within stream and river habitats in order to identify variation between host genotypes, (ii) to compare the microbiota of stone loaches and minnows from stream to those from river to determine possible environmental effects on host-associated microbes and (iii) to compare Rhine and invasive sculpins caught in stream and river in September 2014 with those caught in 2006 to reveal similar patterns of bacterial communities independent from the time point.

Methods

Tissue sampling and DNA extraction

All fish were collected at 10th and 11th of September 2014 at four sites (Figure 3.1). They were caught with hand nets, killed by a blunt blow on the head and wrapped in sterile bags, Whirl-Paks® (Carl Roth). They were snap frozen at -196° C in liquid nitrogen and placed in a super-insulated cryogenic tank for transportation. In the lab, the samples were stored at -20° C for one week before being analyzed. From each fish a third of one pectoral fin was collected under sterile conditions and placed into a 2 ml screw tube containing 50 mg each of 0.1 mm, 0.5 mm and 1 mm glass beads (BioSpec Products) and lysing solution (600 µl buffer RLT mixed with 24 µl 500 mM TCEP). The fins were disrupted using a bead beating tissue homogenizer (Precellys®) for three times 15 s at 6500 g and refrigerat-

ed at -20°C until further processing. DNA was extracted with the AllPrep® 96 DNA/RNA Kit (Qiagen) following the manufacturer's instructions.

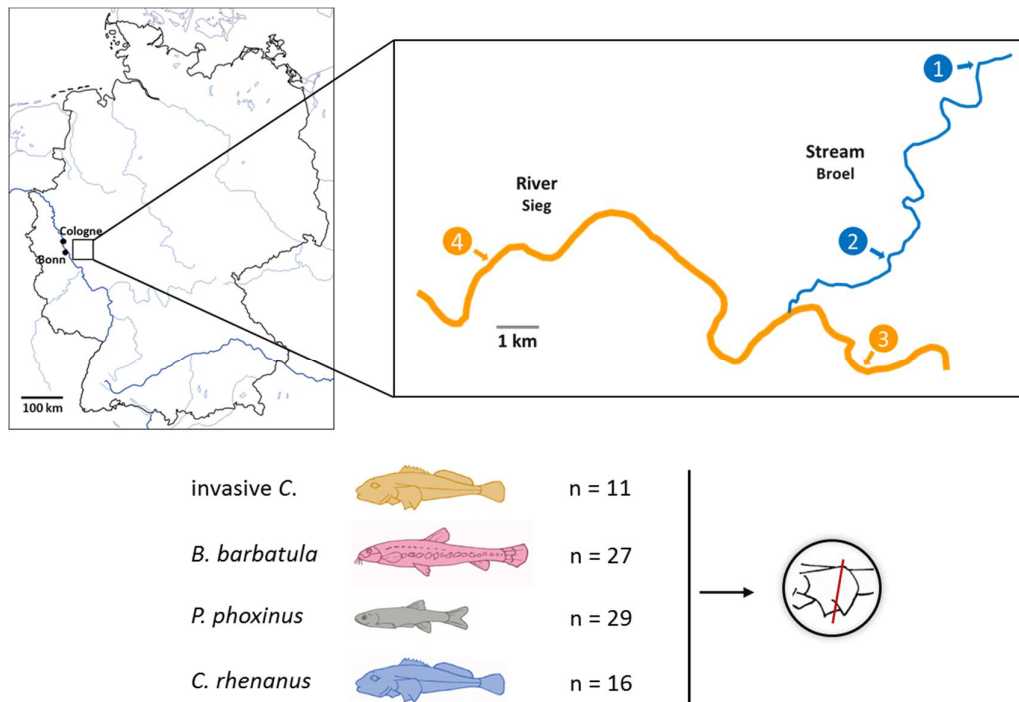


Figure 3.1: Sampling sites along the ecotone Sieg/Broel near Bonn, Germany, where a small stream disemboques into a larger river. Caught were invasive *Cottus* at sites 3 and 4, *B. barbatula* at sites 1-4, *P. phoxinus* at sites 1-4 and *C. rhenanus* at sites 1 and 2. This system enables to study environmental effects on bacterial composition of the same genotype of *B. barbatula* and *P. phoxinus* but not *Cottus* sp. The bacterial communities of each individual fish (of 1/3 of one pectoral fin) were investigated by deep sequencing of a 16S rDNA region and subsequent analysis. Map modified after: <http://www.d-maps.com>

Library preparation

The library was prepared as described in chapter I with some exceptions. Briefly, the 515F – 806R primer pair (Caporaso *et al.* 2011) spanning the V3-4 hypervariable region of the 16S rRNA gene from each fish was used for amplification and individually barcoding following the dual indexing strategy developed by Kozich *et al.* (2013). PCR was carried out using the following conditions: 98°C for 30 s, 98°C for 9 s, 55°C for 30 s, 72°C for 30 s, 72°C for 10 min with 30 cycles, including negative controls for each barcode combination and a positive control per PCR reaction. The mastermix contained 3 μl DNA, 15.25 μl PCR water, 5 μl buffer, 0.5 μl dNTPs, 0.5 μl forward and reverse primer, each, and 0.25 μl Phusion II Hot Start Polymerase. PCR amplicons were separated on a 1.5 % agarose gel; the fragments were visualized and quantified using a Gel Doc™ XR+ System (Bio-Rad) image system. The concentration was assessed for each row of wells separately, based on relative light intensity related to the ruler as internal standard. PCR products were pooled to one subpool per row at equimolar concentrations, respectively, run on a 2 % agarose gel and the amplicons were excised at expected size (~ 250 bp) by using the MiniElute® Gel Extraction Kit (Qi-

agen) according to the manufacturer’s protocol. At the end, the DNA concentration of each subpool was quantified by using Quant-iT™ dsDNA Broad-Range Assay Kit (Life Technologies) on a NanoDrop® ND-3300 Fluorospectrometer to pool them in equimolar concentrations. Then, the generated library was purified with AMPure beads (Agencourt) and finally quantified with Agilent 2100 Bioanalyzer using the Agilent DNA 7500 Kit (Agilent Technologies). The amplicon library was sequenced on the Illumina® MiSeq system using the V2 kit mixed with Illumina generated phiX control libraries.

Sequence processing

Raw sequences were de-multiplexed with a perl script and the paired reads were then quality filtered, trimmed and merged using USEARCH v. 8.0.1616 (Edgar *et al.* 2011). Hereafter, the analysis pipeline within the MOTHUR software v.1.31 (http://www.mothur.org/wiki/MiSeq_SOP) (Schloss 2009) was followed, as described in chapter I. Briefly, Sequences were aligned to SILVA database (Pruesse *et al.* 2007; Quast *et al.* 2013), trimmed to the same length and position and unique sequences were pre-clustered (Schloss *et al.* 2011). UCHIME algorithm (Edgar *et al.* 2011), implemented in MOTHUR served to remove chimeras and subsequently the sequences were classified against the Ribosomal Database Project (RDP) training set v. 9 (Wang *et al.* 2007) with a 80 % bootstrap threshold. With an average neighbor algorithm sequences were then clustered to operational taxonomic units (OTUs) at 97 % similarity level. Finally, the samples were randomly rarified to the smallest read number, 1245, each, to be able to perform and compare diversity measurements at the same sequencing depth. At the end, 72 samples remained for subsequent analysis.

Table 3.1: Sample distribution of genotypes over habitat. Sampling sites were distant from the sculpin hybrid zone so that in the river we only caught invasive *C.* and in the stream only *C. rhenanus*. Hence, there are different genotypes in different environments and no independence.

Genotype	River	Stream
invasive <i>Cottus</i>	11	0
<i>P. phoxinus</i>	7	11
<i>C. rhenanus</i>	0	16
<i>B. barbatula</i>	13	14

Statistical analysis

All statistical tests were performed using a rarified subset of 1245 reads/sample and were done in the R environment (R Core Team 2015), if not stated otherwise. To analyze and compare microbial communities between distinct fish species and habitats I first assessed the relative bacterial taxonomy with nonparametric Kruskal-Wallis rank sum tests and post-hoc pairwise Wilcoxon rank sum tests. By applying several α -diversity tools implemented in MOTHUR (Schloss 2009), I investigated the bacterial species richness (number of observed OTUs), diversity (Shannon’s H) and estimated richness (Chao 1) within each fish individual. The indices were squared, square root- or log transformed if required, to meet the normality requirements, and linear models were applied. The best model was chosen

by using Akaike's information criterion (AIC). To analyze bacterial communities between individuals β -diversity measurements were assessed based on abundance (Bray-Curtis) and presence/absence (Jaccard) of shared OTUs with the 'vegan' package v. 2.3-0 in R (Oksanen *et al.* 2013). The metrics were analyzed with permutational multivariate analysis of variance (ADONIS) and analysis of principal coordinates (PCoA), implemented in 'vegan'. Phylogenetic analyses (weighted and unweighted UniFrac) (Lozupone & Knight 2005; Lozupone *et al.* 2011) were obtained with FastTree v. 2.1.7 (Price *et al.* 2009, 2010) and MOTHUR and subsequently analyzed with ADONIS and PCoA, either. A canonical correspondence analysis (CCA) implemented in 'vegan', directly uses OTU abundance tables and identifies just the variation that can be explained by the applied factors. Afterwards, an ANOVA-like permutation test (`anova.cca`) served to evaluate the significance of the model, the constrained axes and the marginal terms. To identify associations between species, here OTUs, and combinations of sites, here fish individuals, indicator species analyses (Dufrene & Legendre 1997; Caceres & Legendre 2009; De Caceres *et al.* 2012) were performed with the package 'indicspecies' implemented in R based on 10^5 permutations and corrected for multiple testing with false discovery rate (FDR) (Benjamini & Hochberg 1995).

Results

Four distinct fish lineages have been collected at four sampling sites along the contact zone Sieg/Broel near Bonn, Germany, in September 2014. After sequence processing 72 samples remained for further analysis (Table 3.1). The sampling sites were distant from the actual hybrid zone so I assumed that the sculpins in the river were pure invasive *C.* and the sculpins in the stream were pure *C. rhenanus*. In the following I will first demonstrate that distinct fish genotypes in stream or river largely vary in their bacterial composition and will further elaborate general differences between the habitats. Second, I reveal that distinct fish species in the same environment are more different among each other than the same species in different habitats. Lastly, I will illustrate that bacterial microbes of sculpins are in part host-specific independent of varying sampling dates and methods.

Distinct teleost fish contain genotype-specific microbial communities

In order to test whether the bacterial composition in distinct fish genotypes varies I compared the following groups within the river: stone loach (`riv_stl`, $n = 13$), minnow (`riv_min`, $n = 7$) and invasive sculpin (`riv_inv`, $n = 11$), and within the stream: stone loach (`str_stl`, $n = 14$), minnow (`str_min`, $n = 11$) and Rhine sculpin (`str_rhen`, $n = 16$). In sum, I obtained significant differences between the fish species in both habitats, each, for all but one bacterial community structure analysis.



Figure 3.2: Relative frequencies of bacterial classes of teleost fish skin sampled in the contact zone Sieg/Broel in 2014 based on MiSeq analysis. (A) Invasive *Cottus*, *P. Phoxinus* and *B. barbatula* were caught in river Sieg and (B) *C. rhenanus*, *P. Phoxinus* and *B. barbatula* in stream Broel, respectively. Significant differences were obtained between all fish within each habitat as well as among the habitats. Fish in the stream were increased in Alphaproteobacteria and those in the river were enriched in Verrucomicrobiae.

Distinct fish harbor different bacterial taxa

The taxonomic analysis of relative frequencies of bacterial phyla indicated significant variation in three out of six major phyla between all three fish genotypes in stream and river, respectively. Overall, Proteobacteria was the most abundant phylum (river: 30.5%, stream: 49.5%), followed by Bacteroidetes (river: 19.4%, stream: 19.6%), Verrucomicrobia (river: 29%, stream: 11.6%), a proportion of unclassified reads (river: 13.3%, stream: 12.1%) and Firmicutes (river: 2.6%, stream: 2.3%), which together accounted for > 95% of all reads. Proteobacteria mainly consisted of Alpha- (river: 9.4%, stream: 20.0%) and Betaproteobacteria (river: 6.8%, stream: 14.6%) and of less Gammaproteobacteria (river: 5.9%, stream: 7.6%). The second most abundant class was Verrucomicrobiae (river: 26.9%, stream: 9.2%) and Sphingobacteria (river: 9.8%, stream: 7.6%) were also present.

The comparison between the three fish genotypes within each habitat resulted in significant variation in three major phyla (Figure 3.2). In the river the abundance of Proteobacteria in invasive sculpins was heavily decreased compared to the other fish (pairwise Wilcoxon Rank Sum Test riv_rhen:riv_min $P \leq 0.001$, riv_rhen:riv_stl $P \leq 0.001$, riv_min:riv_stl $P = 0.12$). This was also true for Rhine sculpins in the stream (PAIRWISE WILCOXON RANK SUM TEST, str_inv:str_min $P = 0.03$, str_inv:str_stl $P \leq 0.001$, str_min:str_stl $P = 1.00$). According-

ly, all fish contained significantly different abundances of Alpha- Beta- and Gammaproteobacteria (KRUSKAL-WALLIS RANK SUM TESTS, all $P < 0.05$). In both habitats sculpins were enriched in Verrucomicrobia (river: 56.4 %, stream: 24.2 %) and hence in Verrucomicrobiae (river: 54.0 %, stream: 20.6 %). Stone loaches contained a smaller proportion of Verrucomicrobia (river: 18.9 %, stream: 5.2 %) and minnows very little (river: 5.1 %, stream: 1.4 %). Further significant differences at phylum level were obtained for Firmicutes in river (PAIRWISE WILCOXON RANK SUM TEST, riv_rhen:riv_min $P = 0.004$, riv_rhen:riv_stl $P = 0.006$, riv_min:riv_stl $P = 0.006$) and stream (pairwise Wilcoxon Rank Sum Test, str_inv:str_min $P = 1.00$, str_inv:str_stl $P = 0.01$, str_min:str_stl $P = 0.073$). At class level, statistical analyses revealed significant variation in all other major taxa (KRUSKAL-WALLIS RANK SUM TESTS $P < 0.05$ in all cases).

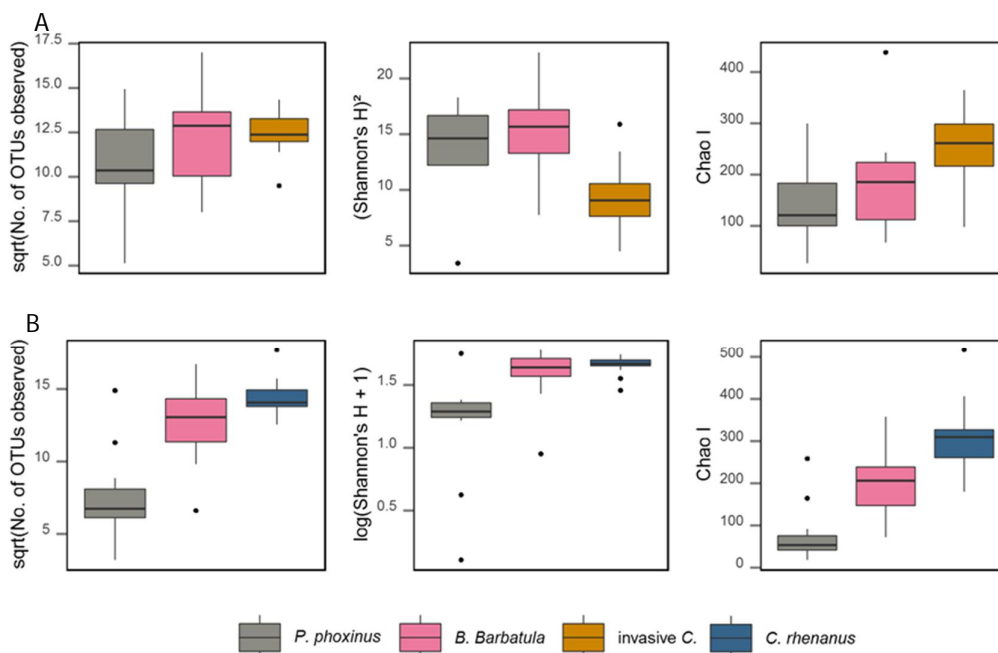


Figure 3.3: α -diversity metrics of bacterial communities within distinct teleost fish skin samples. Fish were caught in the river (A) and in the stream (B). Results were log transformed, square root extracted or squared, if necessary, to meet the normality requirements. Bacterial richness, Shannon diversity and Chao I estimated richness significantly differed between all fish originating from the stream, whereas diversity and estimated richness varied between fish occurring in the river (ANOVA, $P < 0.001$).

Bacterial diversity within and between distinct fish samples was mainly influenced by host species

Several α -diversity measurements aimed at examining bacterial abundance and diversity within distinct fish species (Magurran 1988). I used the following metrics in all subsequent α -diversity analyses: species richness (no. of OTUs), diversity (Shannon's H) and estimated richness (Chao I). Linear models revealed that within both habitats all measurements significantly differed between the three fish species ($P < 0.05$, in all cases), except for species

richness in the river (Table 3.S1-3.S4). The richness and estimated richness was always low in the minnows (Figure 3.3) and high in the sculpins. By contrast, the diversity was increased in the minnows.

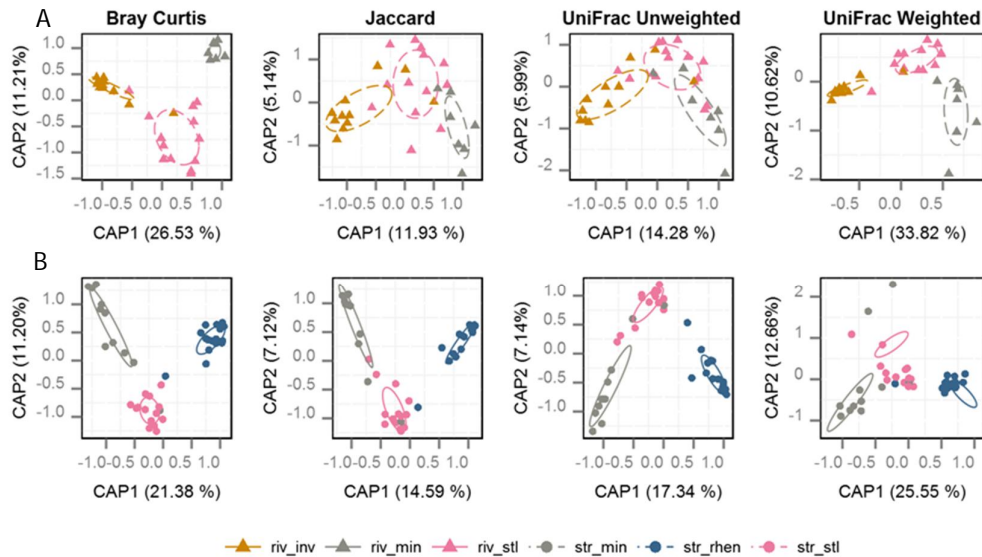


Figure 3.4: Principal coordinate analysis of β -diversity metrics of bacterial communities on different teleosts' skin sampled in two distinct habitats. Examined fish genotypes were invasive *C.* (orange), *B. barbatula* (pink), *P. phoxinus* (grey) and *C. rhenanus* (blue). (A) Individuals in the river are denoted with triangles and those in the stream (B) with circles. Illustrated are the first two axes with the respective variation in every index; ellipses denote the standard deviation of the centroid around the mean of the respective group. The ordination demonstrates a clear separation of all fish within each habitat and reflects the results of the permutational analysis of variance (ADONIS $P \leq 0.001$ for every index).

β -diversity examines differences in bacterial microbiota between individuals and I applied Bray Curtis and Jaccard indices, as well as the phylogenetic metrics Unweighted/Weighted UniFrac in all following β -diversity analyses. Permutational analysis of variance (ADONIS) revealed significant variation between all fish species in both habitats in all metrics (ADONIS, $P \leq 0.001$), which mirrored previous results (Table 3.S5, 3.S6). Subsequent ordination of principal coordinate analysis (PCoA) on all β -diversity measurements clearly illustrated the distinctness of all fish genotypes (Figure 3.4B, C), either. Next, I applied canonical correspondence analysis (CCA) on raw OTU data which tries to display only the data variation which can be explained by the included constrains. As a result, the CCA reflected previous diversity measurements and identified a significant effect of fish genotype on bacterial communities in the river ($F_{1,28} = 2.265$, $X^2 = 1.09$, $P \leq 0.001$), as well as in the stream ($F_{1,38} = 2.304$, $X^2 = 1.262$, $P \leq 0.001$). In both habitats I obtained two significant axes for the best model (river: total inertia: 7.828, explained inertia: 1.09/13.92 %; stream: total inertia: 11.6678, explained inertia: 1.262/10.812 %) (Figure 3.5B, C).

In sum, all analysis of bacterial community abundance and diversity in the distinct fish species indicated the same clear pattern: bacterial composition in stone loaches, minnows and invasive and Rhine sculpins, respectively, inhabiting the contact zone Sieg/Broel is host species-specific. Additionally, I demonstrated that the bacterial composition in both habi-

tats vary which might be caused by environmental effects.

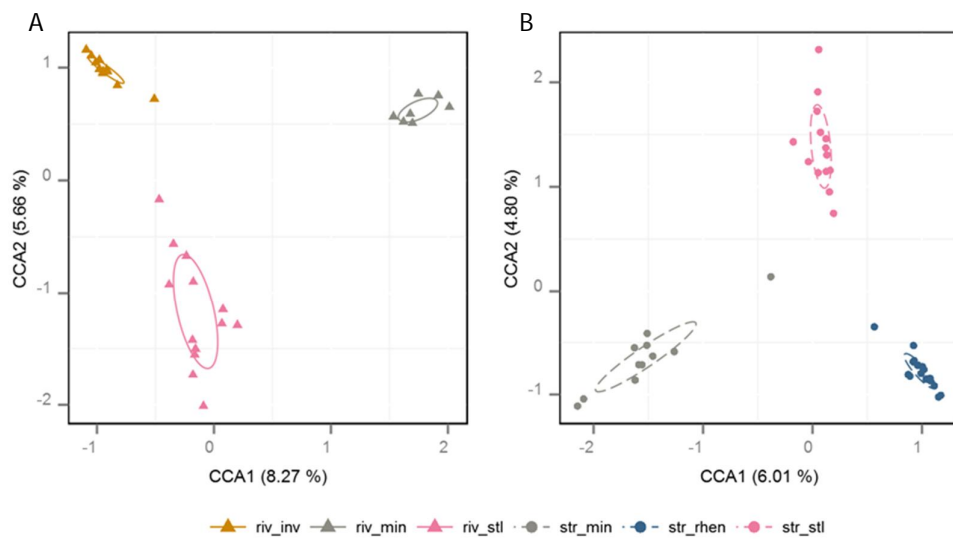


Figure 3.5: Ordination of CCA of distinct teleost fish caught in different habitats. Invasive sculpins (inv), minnows (min) and stone loaches (stlo) were caught in the river (A) and Rhine sculpins (rhen), minnows and stone loaches in the stream (B), respectively. Ellipses denote the standard deviation of the centroid around the mean of the respective group. In both habitats all fish lineages are clearly separated, bacterial communities are influenced by host genotype (permutation test $P \leq 0.001$).

Indicator species analysis revealed a higher abundance of indicative OTUs in the stream and most of them were sculpin-specific

To assess whether certain bacterial OTUs are associated to a particular host or a group of hosts I applied indicator species analyses (Dufrene & Legendre 1997; Caceres & Legendre 2009; De Caceres *et al.* 2012). This resulted in 130 OTUs in the stream fish whereof 88 were associated to Rhine sculpins, 16 to stone loaches, 9 to minnows, 15 to stone loaches and Rhine sculpins, 2 to minnows and stone loaches and none to minnows and Rhine sculpins. The majority of the OTUs belonged to the phyla Proteobacteria, Bacteroidetes and Verrucomicrobia and to the classes Alphaproteobacteria, Verrucomicrobiae and Sphingobacteriia, respectively. In the river fish I obtained only 13 OTUs, 4 associated to invasive sculpins, 2 to stone loaches, 6 to minnow and 1 to invasive sculpins and stone loaches. The frequencies at phylum and class level reflected that of fish in the stream.

Skin bacteria of stone loaches and minnows are rather influenced by host genotype than by host habitat

In order to evaluate whether host-associated bacterial communities in stone loaches and minnows are shaped by environmental or by host genetic effects I compared the two fish species caught at two sampling sites per habitat: stone loach from river (stlo_riv, $n = 13$), stone loach from stream (stlo_str, $n = 14$), minnow from river (min_riv, $n = 7$) and minnow from stream (min_str, $n = 11$). Overall, I revealed larger variation in bacterial composition

between distinct fish species in the same habitat than of the same fish genotype in different habitats, indicating a host-specific bacterial community. I also obtained habitat effects as well as significant interactions between habitat and genotype in some of the analyses.

Taxonomic analyses revealed significant variation between hosts and host habitats

The frequencies of bacterial taxa in minnows and stone loaches varied within each habitat between both species and also between individuals from the same species but distinct habitats (Figure 3.2). I obtained significant differences between minnows and stone loaches for three and two major phyla within the stream and the river, respectively. In both habitats stone loaches contained more Verrucomicrobia than minnows (KRUSKAL-WALLIS RANK SUM TEST stream: $X^2 = 9.1398$, $P = 0.003$; river: $X^2 = 9.7975$, $P = 0.002$). Other varying phyla were Firmicutes and Planctomycetes. Stone loaches and minnows in the stream significantly differed in the abundance of Betaproteobacteria and Sphingobacteriia (KRUSKAL-WALLIS RANK SUM TESTS $P < 0.05$, in all cases). Further, the fish species in both habitats varied in the abundance of Alphaproteobacteria (KRUSKAL-WALLIS RANK SUM TESTS $P < 0.05$, in all cases).

As demonstrated before, fish occurring in the river were generally enriched in Verrucomicrobia, hence I obtained significant differences between the habitats for minnows (KRUSKAL-WALLIS RANK SUM TEST $X^2 = 5.8201$, $P = 0.016$) and stone loaches (KRUSKAL-WALLIS RANK SUM TEST $X^2 = 16.227$, $P \leq 0.001$). The minnows from stream and river further differed in the abundance of Firmicutes (KRUSKAL-WALLIS RANK SUM TEST $X^2 = 6.2094$, $P = 0.013$) and Planctomycetes (KRUSKAL-WALLIS RANK SUM TEST $X^2 = 5.8815$, $P = 0.015$), whereas the stone loaches varied in the amount of Proteobacteria (KRUSKAL-WALLIS RANK SUM TEST $X^2 = 8.1995$, $P = 0.004$) and unclassified reads (KRUSKAL-WALLIS RANK SUM TEST $X^2 = 4.4572$, $P = 0.035$). Alphaproteobacteria (KRUSKAL-WALLIS RANK SUM TEST $X^2 = 14.327$, $P \leq 0.001$) and Betaproteobacteria (KRUSKAL-WALLIS RANK SUM TEST $X^2 = 4.6673$, $P = 0.031$) both accounted for the differences between stream and river in stone loaches and only Alphaproteobacteria (KRUSKAL-WALLIS RANK SUM TEST $X^2 = 5.5505$, $P = 0.018$) significantly differed between minnows from distinct habitats.

Bacterial communities within and between hosts are species-specific but are not independent from habitat effects

By applying several α -diversity measurements I determined the species composition within host samples and analyzed them with linear models. As a result, all metrics revealed a significant effect for genotype ($P < 0.005$, in all cases) but no habitat influence, implying that distinct fish species in the same habitat are more diverse than the same genotype in different habitats. However, there was an interaction between genotype and habitat for bacterial richness, which was higher in the stream than in the river in stone loaches and vice versa in minnows ($P = 0.046$). This was illustrated in Figure 3.3.

Next, I aimed to investigate the bacterial communities between samples and applied numerous β -diversity measurements. Subsequent multivariate statistic indicated a significant influence of genotype and habitat, as well as an interaction between those factors in every

measurement (ADONIS $P < 0.005$).

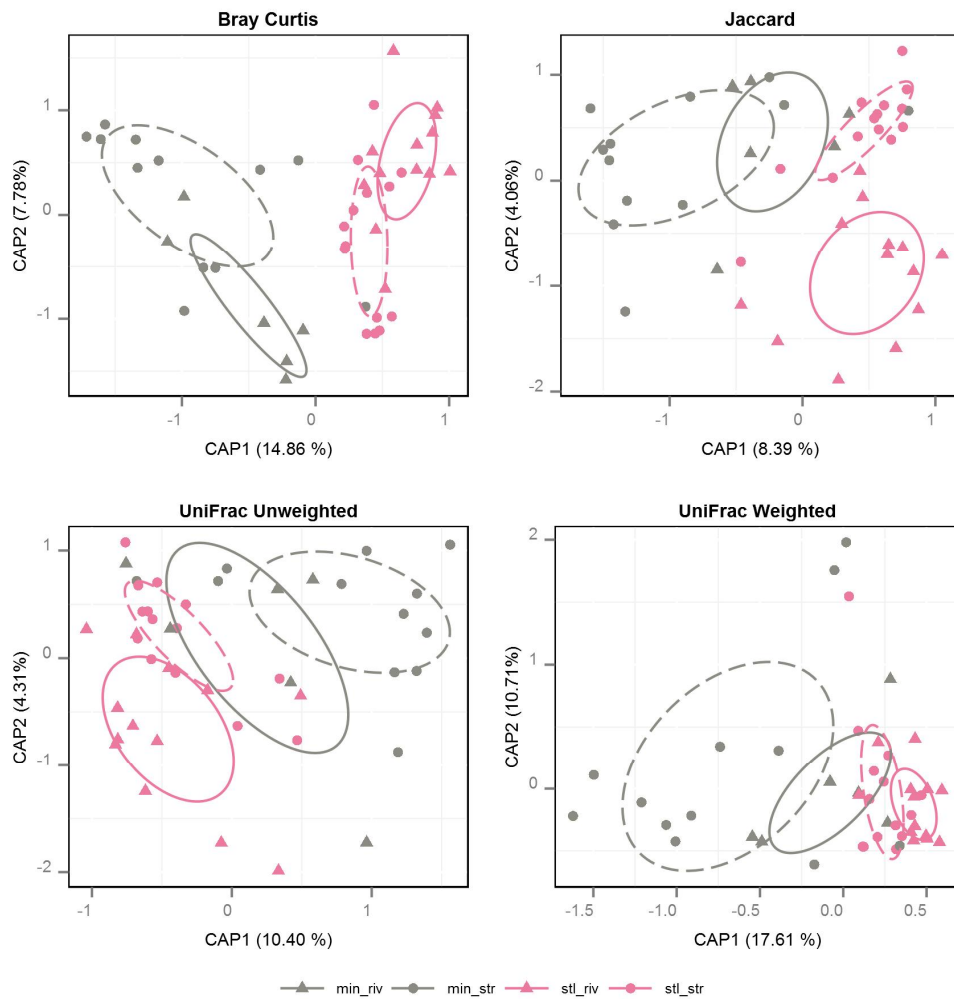


Figure 3.6: Principal coordinate analysis of β -diversity metrics of bacterial communities on different teleost skin sampled in two distinct habitats. Examined fish genotypes were, *B. barbatula* (pink), and *P. phoxinus* (grey). Individuals in the river are denoted with triangles and those in the stream with circles. Illustrated are the first two axes with the respective variation in every index; ellipses denote the standard deviation around the centroid of the respective group. The ordination demonstrates a separation between distinct fish genotypes within each habitat and further between the same fish genotype among both habitats. Thus, it reflects the results of the permutational analyses of variance revealing a significant influence of fish genotype, habitat and an interaction between both factors (ADONIS $P \leq 0.001$ for all terms in every index).

However, genotype always accounted for more variation than habitat, which is supported by ordination of PCoA (Figure 3.4). Thereupon, the fish cluster closer together by genotype than by habitat. A subsequent CCA revealed that genotype ($F_{1,41} = 1.358$, $X^2 = 0.405$, $P \leq 0.001$) and habitat ($F_{1,41} = 1.926$, $X^2 = 0.575$, $P = 0.003$) had a significant influence on bacterial microbiota and the best model (total inertia: 13.568, explained inertia: 1.33/9.81 %) obtained two significant axes (Figure 3.7). Further, there was a significant interaction between genotype and habitat ($F_{1,41} = 1.732$, $X^2 = 0.350$, $P = 0.034$). Overall, the results suggest that minnows and stone loaches living in the contact zone Sieg/Broel exhibit a species-specific bacterial microbiota which is further in part influenced by environmen-

tal variation.

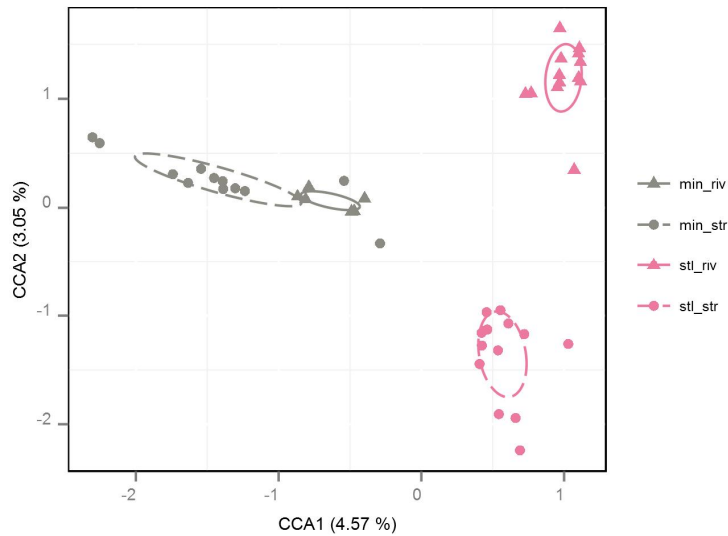


Figure 3.7: Ordination of CCA of minnows and stone loaches at two habitats. Triangles designate fish caught in the river and circles those caught in the stream; minnows are illustrated in grey and stone loaches in pink. Ellipses denote standard deviation around the centroid of the mean of the respective groups. The first axis clearly separates the fish species (but also the minnows from stream and river). The second axis mainly separates the stone loaches by habitat. All included terms, fish genotype and habitat were significant (permutation test $P \leq 0.003$) as well as the interaction between them (permutation test $P = 0.034$).

Distinct sculpin lineages vary in skin bacterial microbiota

In order to investigate if bacterial communities of sculpins recently caught in the contact zone Sieg/Broel are comparable with those caught nine years earlier I examined the following groups: Rhine sculpins from 2005 (rhen_new, $n = 65$) and 2014 (rhen_old, $n = 16$) and invasive sculpins from 2005 (inv_old, $n = 65$) and 2014 (inv_new, $n = 10$). The sculpins from 2005 have been analyzed in Chapter I and only individuals defined as pure genotypes after SNP analysis, exhibiting $> 90\%$ genomic ancestry of invasive or Rhine sculpin, respectively, were included. It should be noted that I only assessed pure fish caught in their original habitat, i.e. Rhine sculpins occurring in stream Broel and invasive sculpins occurring in river Sieg.

I obtained significant variation in bacterial abundance and diversity depending on host genotype within both collections and additionally I identified a clear separation by year. However, some accordance indicates that the two sculpin lineages have specific bacterial microbiota which is constant over time.



Figure 3.8: Relative frequency of bacterial classes of pure sculpins caught in the contact zone Sieg/Broel at two different time points. Fish individuals are grouped by genotype (invasive *C.*, *C. rhenanus*) and collection year (2005, 2014). There are significant differences between the dates and between the same lineages among the dates. Invasive sculpins generally contained more Verrucomicrobia than Rhine sculpins but less Sphingobacteriia and Alphaproteobacteria (KRUSKAL-WALLIS RANK SUM TEST $P < 0.05$, in all cases).

First, I compared the distribution of OTUs between the respective four groups and revealed that the sculpins from 2005 contained and also shared more OTUs, compared to those from 2014. In both collections the Rhine sculpins contained more OTUs and also shared more than invasive sculpins (Figure 3.S1).

The analyses of the frequency distribution of bacterial reads classified to phylum and class level revealed large differences between the dates but also between Rhine and invasive sculpins independent of the sampling year (Figure 3.8). Overall, Proteobacteria was the most abundant phylum (old: 29.0 %, new: 26.5 %) and was in both collections significantly increased in the Rhine sculpins (KRUSKAL-WALLIS RANK SUM TEST old: $X^2 = 21.336$, $P \leq 0.001$; new: $X^2 = 17.778$, $P \leq 0.001$). It mainly consisted of the class Alphaproteobacteria (old: 16.7 %, new: 15.4 %) and less Gamma- (old: 13.9 %, new: 2.2 %) and Betaproteobacteria (old: 6.7 %, new: 7.8 %). Other major phyla were Bacteroidetes (old: 24.6 %, new: 20.5 %), Verrucomicrobia (old: 20.7 %, new: 37.7 %) and Firmicutes (old: 2.9 %, new: 1.1 %). Further major classes were Verrucomicrobiae (old: 18.7 %, new: 34.5 %) and Sphingobacteriia (old: 18.1 %, new: 15.1 %). Rhine and invasive sculpins significantly differed in the abundance of all major phyla in the old collection and in the amount of three major phyla in the new collection (KRUSKAL-WALLIS RANK SUM TESTS $P < 0.05$, in all cases). The fish genotypes also signif-

icantly varied in the abundance of four and three major classes in the old and new collection, respectively (KRUSKAL-WALLIS RANK SUM TESTS $P < 0.05$, in all cases).

In order to determine differences in bacterial diversity between the groups I applied a canonical correspondence analysis (CCA) on raw OTU data. An ANOVA-like permutation test (anova.cca) revealed that the best model (total inertia: 12.238, explained inertia: 0.715/5.84 %) contained the significant terms collection year ($F_{1,152} = 2.967$, $X^2 = 0.225$, $P = 0.001$) and genotype ($F_{1,152} = 4.644$, $X^2 = 0.352$, $P = 0.001$) and an interaction between them ($F_{1,152} = 1.819$, $X^2 = 0.138$, $P = 0.002$). Additionally, I obtained three axes whereof two were significant. The ordination plot illustrates first a clear separation of the sculpin genotypes within the earlier and the later collection and second an explicit separation of the collection dates. However, the two groups of invasive sculpins are close whereas the groups of Rhine sculpins are very distinct (Figure 3.9).

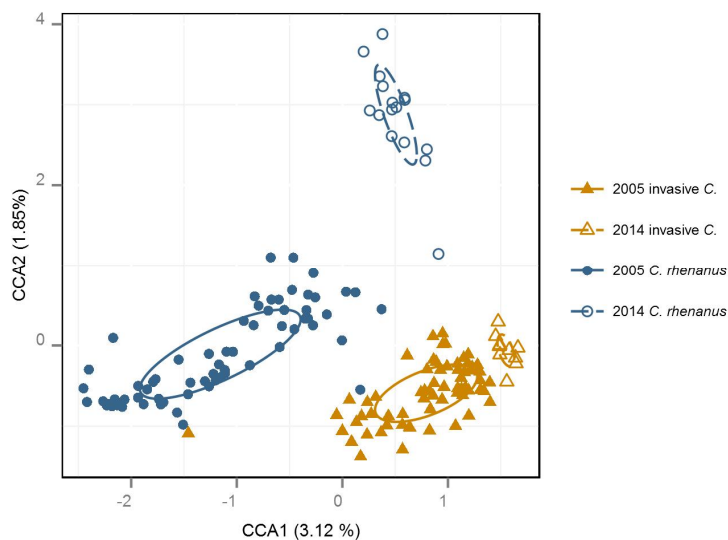


Figure 3.9: Ordination of CCA of pure sculpins caught at different time points. Samples are grouped by sampling years (2005, 2014) and genotype by habitat (invasive *C.* from river, *C. rhenanus* from stream). The earlier samples were investigated in Chapter I and the later samples in the current chapter. Ellipses denote standard deviation around the centroid of the mean of the respective group. The first axis clearly separates the fish genotypes from 2005, whereas the second axis separates the later caught *C. rhenanus* from the rest. The terms date and genotype were significant (permutation test, $P \leq 0.001$).

Discussion

Although it has been well documented that the microbiome influences its host in many aspects, little is known about its role in adaptive speciation processes and in the emergence of new species. To understand these evolutionary dynamics it is fundamental to assess how bacterial communities are shaped and to unravel the impact of host environment and host genetics. Previous work investigated to what extent skin-associated bacteria act as selective force in a secondary contact zone of freshwater sculpins. Deep sequencing of the bacterial 16S rDNA revealed a major influence of host habitat on the microbiota but

host genetic architecture played a role, either. However, in this system the host genotypes were not fully independent from the environment and further experiments were required to disentangle both effects from each other. Accordingly, I examined the contact zone a second time and besides both parental sculpin lineages, Rhine and invasive sculpin, I further included two other fish species which occur in both habitats stream and river without being genetically differentiated. The goal was to assess whether bacterial communities are habitat specific or host genotype specific or both. I revealed whether the effects I obtained for the sculpins can be applied to other fish species inhabiting the contact zone Sieg/Broel, as well.

The analysis of distinct fish genotypes within stream and river, respectively, identified general differences in bacterial community structures within and between both habitats. The results showed a host-specific bacterial microbiota for all particular fish lineages. Bacterial communities of Minnows and stone loaches occurring in stream and river are significantly influenced by host genotype but also habitat. However, fish of the same species but from different habitats are still more similar than distinct fish species inhabiting the same habitat. This suggests that here bacterial communities are fundamentally determined and selected by host species.

The host is influencing its microbiota

Most studies about bacterial microbiota focused on gut (Ley *et al.* 2008; Linnenbrink *et al.* 2013; Ye *et al.* 2014) whereas skin was neglected for a long time. However, skin is, particularly in aquatic organisms, in permanent and direct contact with the environment and acts as first barrier against pathogens (Lazado & Caipang 2014). It provides numerous distinct niches shaped by host physiology and physical conditions like water current. The skin harbors mutualists and commensals but also pathogens so the host has to permanently balance the equilibrium between fighting against pathogens but tolerating harmless bacteria (Gomez *et al.* 2013). Thus, it seems reasonable to hypothesize that the bacterial communities of fish skin is influenced and shaped by host genetics. The fish genotypes I examined in the present study all differed greatly in their bacterial communities whether they occur in the same habitat or not. I determined highly significant differences in bacterial richness and abundance within the lineages in the river and in the stream (Table 3.S1-3.S4) and also between the lineages (verified by permutational analysis of variance on β -diversity measurements (ADONIS) $P \leq 0.001$). Interestingly, the indicator species analysis revealed a relatively high number of significant OTUs in the stream fish ($P < 0.05$), whereof the most are associated to the Rhine sculpins. This was reflected by the ordination plots of β -diversity measurements (Figure 3.4), wherein the sculpins in the stream cluster farther apart than the minnows and the stone loaches. Subsequently, I analyzed minnows and stone loaches caught in both habitats to disentangle host genetic from environmental effects and identified that individuals of the same species but occurring at different habitats were more similar than fish from distinct species but from the same habitat (Figure 3.5). This was supported by linear models applied on α -diversity measurements (Table 3.S9, 3.S10) revealing a significant effect of host genotype ($P \leq 0.001$) but not habitat ($P \geq 0.05$). Multivariate sta-

tistics on β -diversity metrics confirmed the significant influence of fish genotype ($P \leq 0.001$) but also habitat ($P \leq 0.005$) and an interaction between them. However, host genotype accounted for more variation in the data (Table 3.S11) and PCoA graphical ordination (Figure 3.6) demonstrated a closer clustering of fish samples by genotype than by habitat. The first axis of the CCA ordination differentiates minnows from stone loaches and also minnows occurring in the stream from those inhabiting the river; the second axis separates stone loaches from stream and river (Figure 3.7). These results suggest the idea of a species-specific core bacterial microbiota which is relatively stable even under ecological variation caused by changing biotic or abiotic factors. It has been demonstrated in other systems, from hydra (Fraune & Bosch 2010; Franzenburg *et al.* 2013) to humans (Goodrich *et al.* 2014), that closely related individuals have more similar bacterial communities than distantly related ones. There also some examples in fish (Roeselers *et al.* 2011; Larsen *et al.* 2014a; Boutin *et al.* 2014).

The comparison of sculpins caught in 2005 and in 2014 additionally confirmed the idea of host-specific bacterial communities: the ordination plot of CCA illustrated a similar separation by genotype within both collections (Figure 3.7) independent of methodological, seasonal or environmental impacts. However, both fish genotypes are each inhabiting distinct habitats, thus it is not possible to disentangle how much of the variation in bacterial composition can be explained by host genetic or environmental differences.

There are several studies reporting that host genetics shape skin bacterial communities in fish and other species as well. Human skin microbiota has been thoroughly investigated and there is evidence that bacterial communities vary between body parts and individuals (Grice *et al.* 2009). These variations may have been caused by host physiology, diet, immune system, personal habits and local parameters like temperature, pH and humidity (Costello *et al.* 2009; Grice & Segre 2011). Stone loaches and sculpins have several similar life history traits: both species live close to gravel, stones and roots where they hide during the day and search for food during the night (Kottelat & Freyhof 2007). In contrast, minnows show shoaling and schooling behavior, i.e. they are swimming in groups and in coordinated manners (Kottelat & Freyhof 2007). α - and β -diversity measurements didn't reveal a larger accordance in bacterial richness and diversity between sculpins and stone loaches. However, indicator species analysis determined 15 and 1 OTU(s) in individuals caught in the stream and in the river, respectively, which are significantly associated to stone loaches and sculpins but not to minnows. This result confirms the hypothesis that similar lifestyle may account for some agreement in microbial composition in the present study.

Habitat effect

A significant effect of habitat was identified in the taxonomic analysis and also in bacterial richness (Figure 3.2, 3.3). Fish in the stream harbored more Alpha- Beta- and Gamma-proteobacteria and those in the river were enriched in Verrucomicrobiae (Figure 3.2), thus reflecting the results of Chapter I (Figure 3.4). This effect was also confirmed by multivariate statistics, like PCoA, indicating that minnows and stone loaches harbor different bacte-

ria depending on the habitat they occur in (Figure 3.6, Table 3.S11). However, in all analyses habitat accounted for less variation in the data than host genotype (Table 3.S11, 3.S12). The ordination of the CCA revealed a close clustering of the minnows, independent of habitat, but stone loaches from stream and river were very distinct (Figure 3.5). Thus, I conclude that habitat is causing some variation in skin bacterial communities of minnows and stone loaches but host genotype here has the major impact. The habitats stream and river, although very closely located, presumably differ by numerous environmental factors, like temperature, oxygen and water current; therefore, differences in bacterial composition were expected. Indeed, as recorded by the 'Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen' the temperature in stream and river differs during the summer: it reaches at maximum 20°C in the stream and can be > 20°C in the river (<http://luadb.lids.nrw.de/LUA/hygon/pegel.php>). Further, I can't rule out the influence of varying dietary availability in both particular habitats.

It has been hypothesized that bacteria on skin resemble microbes from the respective surrounding by comparing skin samples with aquarium tank biofilm or rocks and logs in nature, respectively (Boutin *et al.* 2013; Fitzpatrick & Allison 2014). However, other studies demonstrated a clear differentiation between microbiota on fish skin and in water or soils (Sullam *et al.* 2012; Larsen *et al.* 2015). Generally, it is questionable to compare such a complex, patchy chemical and physical environment as skin which provides numerous variable niches, with water and sediment samples.

In sum, both examined habitats differ in numerous environmental parameters supporting the results indicating a habitat effect on bacterial communities in fish.

Conclusion

This study aimed at investigating whether skin bacterial communities of teleost fish species are influenced by habitat or host genotype or both. By opposing the same fish species occurring in distinct habitats to different fish species occurring in the same habitat I was able to disentangle genetic from environmental effects. The results confirm a species-specific bacterial microbiota for all of the three examined fish lineages. Apparently, fish are very selective for the microbiota on their skin and not only reflect the bacteria of their environment. However, depending on whether the sampled fish were caught in the stream or in the river I identified some differences in bacterial composition. The results in chapter I, examining two lineages of hybridizing sculpins and their offspring, revealed a major impact of habitat on sculpin-associated bacteria and less effect of host genetics. But it has to be considered that although all sculpin genotypes occurred in either habitat dispersal was limited because there was selection against migrants. Thus, host genetic and environmental effects could not be entirely separated and there was need for fish species inhabiting the whole river system which applied for the minnows and stone loaches. By the current study I could demonstrate that skin bacterial communities are likely influenced and selected by several confounding factors including host genetic architecture being linked to biotic and abiotic variables.

Supplement

Figures

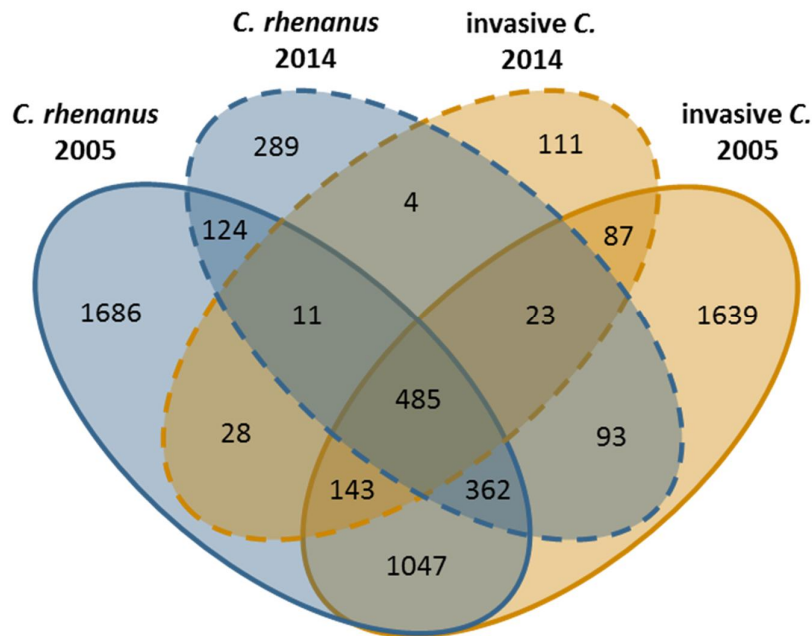


Figure 3.S1: Venn diagram illustrating OTUs shared among the sculpins genotypes (*C. rhenanus*, invasive *C.*) caught at different dates (2004 and 2014) and used in chapter I and II, respectively. Sculpins caught in 2014 contained less OTUs than those of 2014. This was caused by lower read quality. The amount of OTU shared between the lineages is similar.

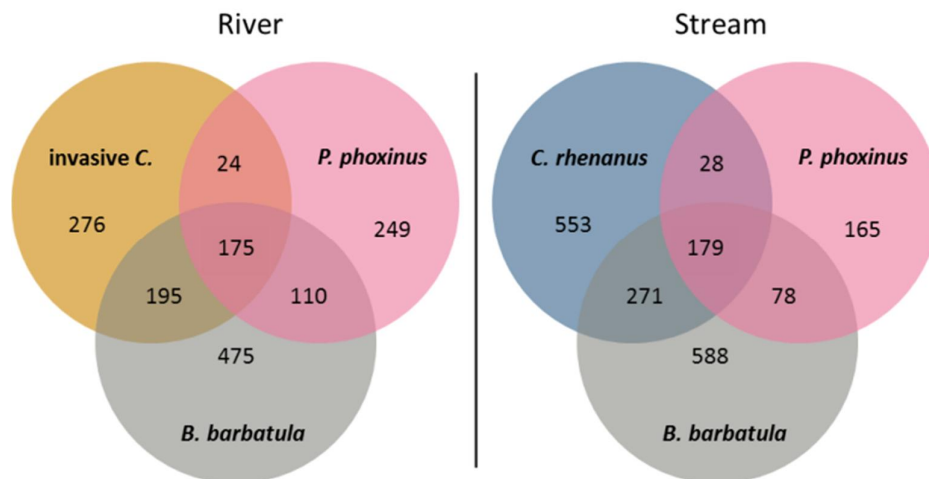


Figure 3.S2: Venn diagram illustrating OTUs shared between host genotypes among the habitats river and stream. Fish in the stream have more OTUs in total than those in the river. The minnows have least OTUs and the stone loaches the most.

Tables

Table 3.S1: Linear model of bacterial α -diversity of fish in the river with significance of model terms. The full model contained the factor genotype (G): *P. phoxinus* (min), invasive *C. (inv)*, *B. barbata* (stlo) and is also the best model, selected by minimizing the Akaike Information Criterion (AIC). The fish genotypes are significantly different in Chao I estimated richness and Shannon's diversity.

Index	Term	Df	Sum Sq	Mean Sq	F value	Pr (>F)
sqrt(Species richness)	G	2	13.844	6.9221	1.2843	0.2926
	Residuals	28	150.911	5.3897		
Chao I	G	2	52528	26263.9	3.4607	0.04539
	Residuals	28	212500	7589.3		
(Shannon) ²	G	2	211.53	105.766	6.1474	0.006119
	Residuals	28	481.74	17.205		

Bold P-values indicate significance ($P < 0.05$).

Table 3.S2: Linear model of bacterial α -diversity of fish in the river with significance of coefficients. The intercept indicates invasive sculpins. Minnows and invasive sculpins are significantly different in Chao I and Shannon's diversity and stone loaches significantly vary in Chao I.

Index	Term	Estimate	Std. Error	t value	Pr(> t)	adj. R ²
sqrt(Species richness)	(Intercept)	12.4478	0.7	17.783	<2e-16	0.0186
	G (min)	-1.7303	1.1225	-1.542	0.134	
	G (stlo)	-0.3022	0.9511	-0.318	0.753	
Chao I	(Intercept)	251.43	26.27	9.572	2.51e-10	0.1409
	G (min)	-106.45	42.12	-2.527	0.0174	
	G (stlo)	-64.57	35.69	-1.809	0.0812	
(Shannon) ²	(Intercept)	9.323	1.251	7.454	4.05e-08	0.2555
	G (min)	4.123	2.005	2.056	0.04924	
	G (stlo)	5.892	1.699	3.467	0.00172	

Bold P-values indicate significance ($P < 0.05$).

Table 3.S3: Linear model of bacterial α -diversity of fish in the stream with significance of model terms. The full model contained the factor genotype (G): minn, rhen, stl and is also the best model, selected by minimizing the Akaike Information Criterion (AIC). The fish genotypes in the stream are significantly different in all indices.

Index	Term	Df	Sum Sq	Mean Sq	F value	Pr (>F)
sqrt(Species richness)	G	2	313.09	156.547	28.957	2.29e-08
	Residuals	38	205.43	5.406		
Chao I	G	2	343191	171595	29.345	1.97e-08
	Residuals	38	222208	5848		
log(Shannon)	G	2	2.2143	1.10715	13.077	4.77e-05
	Residuals	38	3.2171	0.08466		

Bold P-values indicate significance ($P < 0.05$).

Table 3.S4: Linear model of bacterial α -diversity of fish in the stream with significance of coefficients. The Included term is explained in Table 3.S3, the intercept indicates minnows. The coefficients are significant in every index.

Index	Term	Estimate	Std. Error	t value	Pr(> t)	adj. R ²
sqrt(Species richness)	(Intercept)	7.6124	0.701	10.859	3.29e-13	0.583
	G (rhen)	6.8459	0.9107	7.517	4.96e-09	
	G (stlo)	4.9646	0.9368	5.3	5.19e-06	
Chao I	(Intercept)	79.16	23.06	3.433	0.001454	0.586
	G (rhen)	228.42	29.95	7.626	3.55e-09	
	G (stlo)	117.09	30.81	3.8	0.000508	
log(Shannon)	(Intercept)	1.50771	0.08773	17.186	< 2e-16	0.3765
	G (rhen)	0.55447	0.11396	4.865	2.02e-05	
	G (stlo)	0.47959	0.11723	4.091	0.000215	

Bold P-values indicate significance ($P < 0.05$).

Table 3.S5: Permutation multivariate analysis of dissimilarity matrices for fish in the river. The included term is explained in Table 3.S1 and is significant for every measurement.

Index	Terms	Df	SumsSqs	MeanSqs	F. Model	R ²	Pr (>F)
Bray Curtis	G	2	3.2011	1.60055	6.5303	0.31808	0.001
	Residuals	28	6.8626	0.24509		0.68192	
	Total	30	10.0637			1	
Jaccard	G	2	1.5858	0.79288	2.3054	0.14139	0.001
	Residuals	28	9.63	0.34393		0.85861	
	Total	30	11.2158			1	
Unifrac unweighted	G	2	1.4187	0.70937	2.426	0.14769	0.001
	Residuals	28	8.1873	0.2924		0.85231	
	Total	30	9.606			1	
Unifrac weighted	G	2	1.1085	0.55424	8.3547	0.37373	0.001
	Residuals	28	1.8575	0.06634		0.62627	
	Total	30	2.9659			1	

Bold P-values indicate significance ($P < 0.05$).

Table 3.S6: Permutation multivariate analysis of dissimilarity matrices for fish in the stream. Included term is genotype, which is significant for every measurement.

Index	Terms	Df	SumsSqs	MeanSqs	F. Model	R ²	Pr (>F)
Bray Curtis	G	2	4.016	2.00801	7.4309	0.28114	0.001
	Residuals	28	10.268	0.27022		0.71886	
	Total	30	14.285			1.00000	
Jaccard	G	2	2.9254	1.46272	4.4353	0.18926	0.001
	Residuals	28	12.5321	0.32979		0.81074	
	Total	30	15.4575			1.00000	
Unifrac unweighted	G	2	2.8637	1.4319	5.1525	0.21333	0.001
	Residuals	28	10.5602	0.2779		0.78667	
	Total	30	13.4239			1.00000	
Unifrac weighted	G	2	1.3750	0.68749	7.6451	0.28692	0.001
	Residuals	28	3.4172	0.08993		0.71308	
	Total	30	4.7922			1.00000	

Bold P-values indicate significance ($P < 0.05$).

Table 3.S7: CCA of river samples. The included term is explained in Table 3. S1. Significance of model, terms and axes was evaluated by using permutation tests.

	Df	ChiSquare	F	Pr(>F)
Significance of Model				
G	2	1.0900	2.2648	0.001
Residual	28	6.7376		
Significance of Model terms				
G	2	1.0900	2.2648	0.001
Residual	28	6.7376		
Significance of axes				
CCA1	1	0.6471	2.6891	0.001
CCA2	1	0.4429	1.8405	0.001
Residuals	18	6.7376		

Bold P-values indicate significance ($P < 0.05$).

Table 3.S8: CCA of stream samples. The included term is explained in Table 3. S3. Significance of model, terms and axes was evaluated by using permutation tests.

	Df	ChiSquare	F	Pr(>F)
Significance of Model				
G	2	1.2618	2.304	0.001
Residual	38	10.4055		
Significance of Model terms				
G	2	1.2618	2.304	0.001
Residual	38	10.4055		
Significance of axes				
CCA1	1	0.7015	2.5618	0.001
CCA2	1	0.5603	2.0462	0.001
Residuals	38	10.4055		

Bold P-values indicate significance ($P < 0.05$).

Table 3.S9: Linear model of bacterial α -diversity of minnows and stone loaches across distinct habitats with significance of model terms. The full model contained the factors genotype (G): minn, stlo, and habitat (H): stream, river. The best model was selected by minimizing the Akaike Information Criterion (AIC). For species richness fish genotype, as well as an interaction between genotype and habitat, was significant. For the other indices genotype was significantly explaining the data variation.

Index	Term	Df	Sum Sq	Mean Sq	F value	Pr (>F)
sqrt(Species richness)	G	1	136.06	136.057	17.5608	0.000144
	H	1	9.77	9.768	1.2607	0.268042
	G:H	1	32.73	32.732	4.2247	0.046247
	Residuals	41	317.66	7.748		
log(Chao)	G	1	6.8356	6.8356	18.016	0.000115
	Residuals	43	16.3149	0.3794		
(Shannon) ²	G	1	416.17	416.17	13.547	0.000644
	Residuals	43	1320.95	30.72		

Bold P-values indicate significance ($P < 0.05$).

Table 3.S10: Linear model of bacterial α -diversity of minnows and stone loaches across distinct habitats with significance of coefficients. Included terms are explained in Table 3.S9. The intercept indicates minnows in the stream.

Index	Term	Estimate	Std. Error	t value	Pr(> t)	adj. R ²
sqrt(Species richness)	(Intercept)	7.6124	0.8393	9.07	2.37e-11	0.313
	G (stlo)	4.9646	1.1215	4.427	6.93e-05	
	H (river)	3.1051	1.3458	2.307	0.0262	
	G (stlo): H (river)	-3.5366	1.7206	-2.055	0.0462	
log(Chao)	(Intercept)	4.3635	0.1452	30.055	< 2e-16	0.2789
	G (stlo)	0.7956	0.1874	4.245	0.000115	
(Shannon) ²	(Intercept)	9.788	1.306	7.492	2.51e-09	0.2219
	G (stlo)	6.208	1.687	3.681	0.000644	

Bold P-values indicate significance ($P < 0.05$).

Table 3.S11: Permutation multivariate analysis of dissimilarity matrices of minnows and stone loaches across distinct habitats. The included terms are explained in Table 3.S9. All terms and interactions between them are significant in every measurement.

Index	Terms	Df	SumsSqs	MeanSqs	F. Model	R ²	Pr (>F)
Bray Curtis	G	1	1.9351	1.93507	5.9613	0.11574	0.001
	H	1	0.7843	0.78432	2.4162	0.04691	0.005
	G:H	1	0.6903	0.69033	2.1267	0.04129	0.002
	Residuals	41	13.3087	0.3246		0.79605	
	Total	44	16.7185			1	
Jaccard	G	1	1.0512	1.05123	2.7843	0.05914	0.001
	H	1	0.6688	0.6688	1.7714	0.03762	0.001
	G:H	1	0.5763	0.57629	1.5264	0.03242	0.005
	Residuals	41	15.4799	0.37756		0.87082	
	Total	44	17.7762			1	
Unifrac unweighted	G	1	0.9631	0.96313	3.0232	0.06391	0.001
	H	1	0.5715	0.57154	1.7941	0.03792	0.004
	G:H	1	0.4749	0.47493	1.4908	0.03151	0.019
	Residuals	41	13.0616	0.31857		0.86666	
	Total	44	15.0712			1	
Unifrac weighted	G	1	0.5208	0.52082	5.7579	0.11057	0.001
	H	1	0.2473	0.24729	2.7339	0.0525	0.002
	G:H	1	0.2334	0.23344	2.5808	0.04956	0.002
	Residuals	41	3.7085	0.09045		0.78736	
	Total	44	4.7101			1	

Bold P-values indicate significance ($P < 0.05$).

Table 3.S12: CCA of minnows and stone loaches across distinct habitats. The included terms are explained in Table 3.S9. Significance of model, terms and axes was evaluated by using permutation tests.

	Df	ChiSquare	F	Pr(>F)
Significance of Model				
Model	3	1.3304	1.4857	0.001
Residual	41	12.2378		
Significance of Model terms				
G	1	0.5908	1.9792	0.001
H	1	0.3894	1.3047	0.003
G:H	1	0.3502	1.1732	0.034
Residual	41	12.2378		
Significance of axes				
CCA1	1	0.6207	2.0794	0.001
CCA2	1	0.4144	1.3883	0.001
CCA3	1	0.2953	0.9895	0.469
Residuals	41	12.2378		

Bold P-values indicate significance ($P < 0.05$).

Table 3.S13: CCA of Rhine and invasive sculpins of Chapter I and Chapter II. The included terms are explained in table. Significance of model, terms and axes was evaluated by using permutation tests.

	Df	ChiSquare	F	Pr(>F)
Significance of Model				
Model	3	0.7149	3.1432	0.001
Residual	15	11.5235		
Significance of Model terms				
D	1	0.2249	2.9669	0.001
G	1	0.3521	4.6440	0.001
D:G	1	0.1379	1.8186	0.002
Residual	15	11.5235		
Significance of axes				
CCA1	1	0.3824	5.0447	0.001
CCA2	1	0.2261	2.9823	0.001
CCA3	1	0.1063	1.4025	0.094
Residuals	15	11.5235		

Bold P-values indicate significance ($P < 0.05$).

CHAPTER 3

Makroparasites of distinct *Cottus* lineages across natural hybrid zones

Introduction

Since the 1990s there is growing interest to understand the factors initiating and driving ecologically-based divergent selection. (Schluter 2001, 2009; Rundle & Nosil 2005; Hendry *et al.* 2007; Nosil 2012). During this process, referred to as ecological speciation, barriers to reproduction between populations emerge due to adaptation to different environments or niches. Most studies about ecological speciation focused on biotic habitat differences or sexual selection (reviewed by Schluter 2001; Rundle & Nosil 2005; Lowry *et al.* 2008a; Maan & Seehausen 2011). However, host-parasite interaction, despite representing a major selective force in evolution (Haldane 1949), has gained little attention in speciation research for a long time (Karvonen & Seehausen 2012). Parasites exploit their host in numerous ways and can mediate their behavior, immune system and fitness; hence, they have a fundamental impact and can induce strong selection pressure (Price *et al.* 1986; Wood *et al.* 2007).

If organisms are adapted to their local environment they are expected to have a fitness advantage compared to non-local habitats, this is referred to as local adaptation (Lively & Dybdahl 2000). For instance, a host is less susceptible to a common parasite since it acquired an adaptive immune response against it but not to an uncommon parasite it is not regularly exposed to (e.g. Kaltz & Shykoff 1998; Lively & Dybdahl 2000; Kawecki & Ebert 2004). Otherwise, a parasite is more infective to common hosts than to uncommon hosts. The occurrence of a parasite is depending on various ecological factors, like temperature and the abundance of intermediate and final hosts (Arneberg *et al.* 1998; Karvonen *et al.* 2013a). Thus, host populations in different environments are likely to become reproductively isolated by experiencing different forms and abundances of parasitism causing genetic differentiation in the host and/or different adaptive immune defense mechanisms (Eizaguirre *et al.* 2009). Parasite-mediated selection against immigrants or hybrids, experiencing a higher parasite infection in foreign habitats, can emerge. There are well known studies in fish investigating differences in parasite composition in relation to parasite-driven speciation. It has for instance been demonstrated that two distinct ecotypes of the threespine stickleback (*Gasterosteus aculeatus*) vary in susceptibility to the eye fluke *Diplostomum pseudospathaceum*: the infection rate was higher in river populations, where the parasite has low prevalence, than in those from lakes, wherein the parasite prevalence is high (Scharsack & Kalbe 2014). Further, transplant experiments of marine sticklebacks to freshwater and saltwater resulted in higher parasite infection rates in fish exposed to freshwater and direct growth costs (MacColl & Chapman 2010). Simultaneously, freshwater fish were less susceptible to parasites in their native environment. The authors concluded that in this system parasites contribute to selection against migrants.

A good scenario to study the influence of parasites to speciation processes are naturally emerged hybrid zones where distinct species/populations hybridize and produce many generations of recombinants. New genetic variation arises and selection of specific genotypes can drive adaptive speciation processes. Dispersal can be limited by intrinsic barriers but also by ecological-driven selection, i.e. parasites, the migrants are not adapted to. As a consequence, the genetic analysis of hybrid zones can reveal associations of genotypes with specific genetic backgrounds or environmental factors that highlight the role of genomic regions in evolutionary processes. Several studies investigated the role of parasites in hybrid speciation (Sage *et al.* 1986; Le Brun *et al.* 1992; Wolinska *et al.* 2008; Baird *et al.* 2012; Guttel & Ben-Ami 2014). However, the results were inconsistent and unpredictable. In some studies the hybrids were more resistant (Le Brun *et al.* 1992) than their parents whereas they were more susceptible in others (Moullia 1999; Šimková *et al.* 2013) and in some they were intermediate (Wiley *et al.* 2009). Moreover, we are lacking information about the role of parasites in early speciation processes (Karvonen & Seehausen 2012).

In this study we are using European freshwater sculpins (*Cottus* spp.) to understand the ecological and genetic basis of diverging speciation processes. Sculpins typically live in oxygen-rich and relatively cold and small tributaries throughout Europe. Presumably 200 years ago man-made canals connected the river systems Rhine and Scheldt in Belgium and thereby caused the hybridization of two sculpins species: *Cottus rhenanus*, inhabiting the Rhine system, and *Cottus perifretum*, inhabiting the Scheldt system. In contrast to their parental species, whose habitats are small, cold and oxygen-rich streams, the resulting hybrids invaded larger streams that are relatively warm during the summer and have been free of sculpins before. Thus we refer to them as invasive *Cottus*, as they invaded new habitats. In the 1980s they have for the first time been reported in the Netherlands (De Nie 1997; Nolte *et al.* 2005) and subsequently dispersed upstream the Rhine system with 4-8 km per year (Nolte *et al.* 2005). This resulted in the emergence of secondary contact zones with one of their parents, *C. rhenanus*, where small streams disemboque into larger rivers, near Bonn, Germany (Nolte *et al.* 2005, 2009). In-depth analyses revealed that these contact zones are only about 2 km wide and that Rhine and invasive sculpins hybridize under absence of intrinsic barriers (Nolte *et al.* 2006, 2009). However, selection favors resident sculpins in their habitats, there's selection against immigrants and thus it has been discussed that ecological factors play a key role in shaping these hybrid zones (Nolte *et al.* 2006, 2009). It has been revealed before that host-associated bacterial microbiota in the sculpin secondary contact zone is mainly influenced by host habitat but a host genotype effect is also present (Chapter I, II). We were here particularly interested in comparing the parasite communities and parasite infectivity between different sculpin populations and species in Western Europe. To date, few studies about parasites in *Cottus* spp. exist (Harris 1985; Moravec 2001; Chubb *et al.* 2006; Winger *et al.* 2008; Ieshko *et al.* 2013) but none examined parasite communities and infection rates in Western European sculpins and in relation to parasite-mediated selection. Given that parasites can be a strong selective force for their hosts we hypothesized that sculpins from different populations and lineages

would have adapted their immune system and optimized defense mechanisms against different local parasites.

This study aimed at investigating the influence of different host populations and species, inhabiting or originating from distinct streams and rivers, on parasite abundance and diversity. Occurring variation would indicate different selective pressures caused by different parasite communities which may lead to the emergence of reproductive barriers and adaptive speciation processes.

First, we performed a pilot study to examine the parasite composition of different sculpin populations, representing two different genetic lineages, collected at two distinct streams in the contact zone near Bonn, Germany. Remarkably, one of those streams, Wahnbach, is one of the two only known streams which have partially been colonized by invasive sculpins (Nolte *et al.* 2006). The Wahnbach is retained through a dam and expands to a barrier lake up to a water dam and behind disembogues into the larger river Sieg, typically inhabited by invasive sculpins. Upstream, Wahnbach is colonized by Rhine sculpins whose dispersal is prohibited by the dam. Thus, downstream, previously free of sculpins (Nolte *et al.* 2006), the lowest reaches of the stream have been colonized by invasive sculpins originating from the Sieg (Figure 4.1). Here, the goal was first to make an inventory of parasites present in *Cottus* spp. in Western Europe and second to investigate whether distinct host populations and genotypes experiencing different environments differ in parasite abundance and diversity.

Second, we investigated infection rates of several sculpin populations originating from distinct habitats and locations by exposing them to the frequent fish parasite *Diplostomum pseudospathaceum*. We expected to detect variation in susceptibility depending on host population. Cercariae of *D. pseudospathaceum* develop in snails and are then released into the water to infect fish. In the fish the cercariae migrate to the host's eye lenses within 24 h and can induce the formation of cataracts (Seppälä *et al.* 2011), which reduces the fish's visual abilities and thereby increases the probability to be eaten by a fish-eating bird, the final host of *D. pseudospathaceum* (Niewiadomska 1984). In addition, the parasites are known to alter their host's behavior: mature metacercariae influence the depth preference of the fish so that they swim near the surface and thus can be easily caught by birds (Crowden & Broom 1980). The aim of this study was to test whether distinct host populations differ in susceptibility to *D. pseudospathaceum*.

Methods

Parasite screen of wild caught sculpins

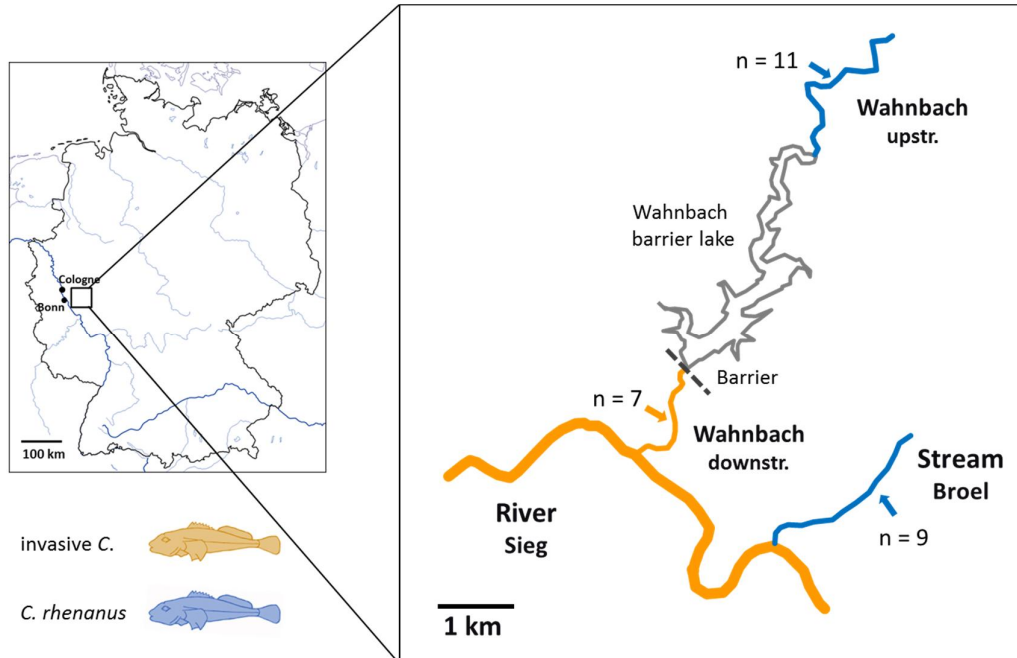


Figure 4.1: Sampling sites near the contact zone Sieg/Broel, Bonn, Germany. Invasive *C.* ($n = 7$) were caught in lower Wahnbach and *C. rhenanus* were caught in Broel ($n = 9$) and upper Wahnbach ($n = 11$). Fish were killed and dissected to examine parasite infections in all external surfaces and inner organs. Map modified after <http://d-maps.com>.

For the present study we caught sculpins at three sites near the area of secondary contact along the hybrid zone Sieg/Broel near Bonn, Germany, in March 2013. The streams Broel and upper Wahnbach are inhabited by Rhine sculpins (Nolte *et al.* 2006) but lower Wahnbach, wherein resident sculpins were absent in 1995 (Freyhof, unpublished), is inhabited by invasive sculpins (Nolte *et al.* 2006). These invasive fish presumably originated from river Sieg (Nolte *et al.* 2006). The Wahnbach is split by a water dam retaining a large water reservoir which acts as barrier for fish so that the sculpin lineages are not able to exceed their range and remain separated (Figure 4.1). In total 27 fish were caught with a hand net, *C. rhenanus* at Broel ($n = 9$), *C. rhenanus* at upper Wahnbach ($n = 11$) and invasive *Cottus* at lower Wahnbach ($n = 7$) and thereafter transported alive to the lab in appropriate containers under constant aeration. Subsequently, they were killed with MS222 and their external surfaces were screened for parasites. Total length (from tip of the snout to the tip of the longer lobe of the caudal fin) and standard length (from the top of the snout to the posterior end, excluding the caudal fin) were measured to the nearest mm and weight to the nearest 0.1 mg (<http://www.fishbase.org/Glossary> accessed 06/08/2016). The fish were decapitated and afterwards the body cavity was opened, spleen and liver were removed and weighted (to the nearest 0.1 mg) and together with all other inner organs, including eyes and gills, examined for parasite abundance under a binocular microscope.

Experimental infection of lab bred sculpins

As previously described, wild *Cottus* spp. can be infected with the eye fluke *Diplostomum pseudospathaceum*. To assess whether different lineages of sculpins vary in their susceptibility they were experimentally exposed to that parasite. We performed two pilot studies with sculpins from different populations to assess if they generally differ in their parasite susceptibility and if it's reasonable to plan further experiments with larger sample sizes. In the first experiment in total 14 lab bred fish were used, nine originated from German populations, (5 x stream Broel, 4 x stream Naaf), four *C. perifretum* from Belgian stream Laarse Beek, and two F1 originated from crosses between *C. rhenanus* (Wahnbach) and *C. perifretum* (Laarse Beek). In the second experiment we included 52 lab bred fish in total, 24 invasive *C.* from two Pleisbach families, 16 *C. perifretum* from two Laarse Beek families and 12 *C. perifretum* from one Witte Nete family (Figure 4.2).

The experiments were carried out as described in Kalbe & Kurtz (2006) with small modifications. Briefly, five naturally infected snails, *Lymnaea stagnalis*, were collected at the lake 'Kleiner Plöner See' and subsequently kept in the lab under natural conditions. They were placed in 50 ml tap water containing glass beakers and were brightly illuminated for 2 h to obtain *D. pseudospathaceum* cercariae. Subsequently, the cercariae of all snails were pooled and counted within additional 2 h, so that they all had the same age (2-4 h). Then, 100 cercariae were pipetted to each fish which were kept in separate tanks containing 10 L water for 24 h. Afterwards, they were killed with an overdose MS222 (1 g/L). They were weighted (to the nearest 0.1 mg) and measured (to the nearest mm) and the eyes were removed and checked for eye fluke presence under a binocular microscope.

Table 4.1: Sample distribution in the two infection experiments. The first experiment included two families of Rhine sculpins, a Scheldt sculpins family and two F1 crosses between Rhine and Scheldt sculpins. The second experiment included two families of invasive sculpins and three families of Scheldt sculpins.

exp.	Origin	Genotype	N
1	Broel	<i>C. rhenanus</i>	5
1	Naaf	<i>C. rhenanus</i>	4
1	Laarse Beek	<i>C. perifretum</i>	4
1	Wahnbach x Laarse Beek	F1 (<i>C. rhenanus</i> x <i>C. perifretum</i>)	2
2	Pleisbach	invasive <i>C.</i> family 14	12
2	Pleisbach	invasive <i>C.</i> family 21	12
2	Laarse Beek	<i>C. perifretum</i> family 2	8
2	Laarse Beek	<i>C. perifretum</i> family 3	8
2	Witte Nete	<i>C. perifretum</i>	12

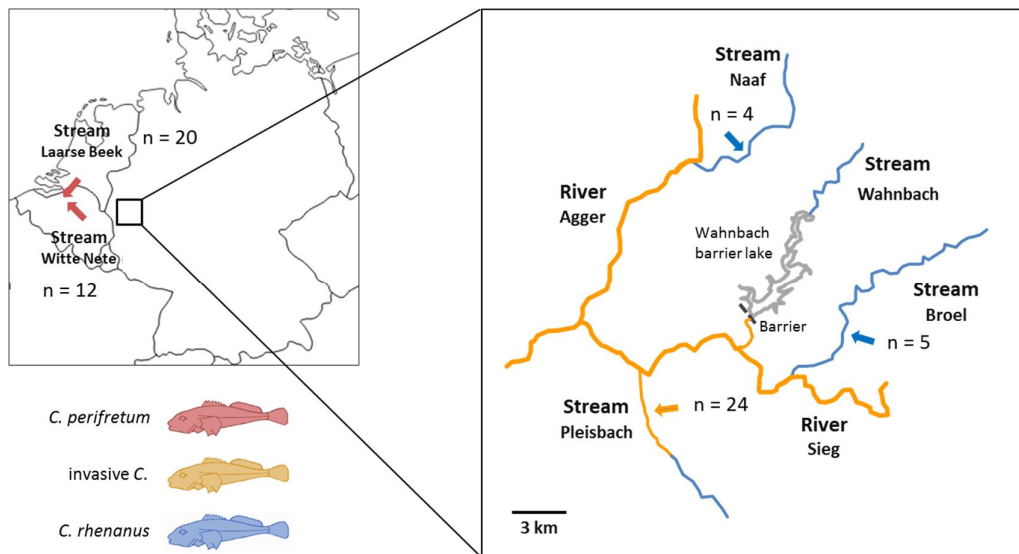


Figure 4.2: Origin of fish populations exposed to *D. pseudospathaceum*. Fish originated from stream Laarse Beek and Witte Nete in Belgium (*C. perifretum*) and from the contact zones near Bonn, Germany. Therein, *C. rhenanus* were confined to small streams and only 30 years ago invasive *C.* colonized larger rivers. The colors refer to the respective fish lineages inhabiting the rivers/streams. Map modified after <http://d-maps.com>

Statistical analysis

Wild caught fish

The wild caught fish were analyzed for parasite prevalence and abundance, following the definitions of Bush *et al.* (1997). Parasites belonging to the class ciliate are usually very abundant and were estimated to be 0, 10, 50 or 100. These taxa can reproduce on the fish and the examined numbers may not account for actual infection rates. Thus, only prevalence (proportion of infected hosts among all the hosts examined) was applied. We didn't identify any influence of host sex and organ weight on parasite abundance/diversity and excluded the factors from the analyses. Body size of the fish was defined by measurements of standard length. The nonparametric WILCOXON RANK SUM TEST was applied on parasite abundance to analyze differences between habitats. For the analysis of variance (ANOVA) we then checked for normal distribution of the residuals and log transformed the data if necessary to meet the normality requirements. Possible influencing factors were determined and the best model was chosen based on Akaike's information criterion (AIC).

Infection experiments

The lab raised fish from the infection experiments were exclusively examined for abundance of *D. pseudospathaceum*. In the first experiment we excluded the F1 fish from further statistical analysis because of low sample size ($n = 2$). By applying WILCOXON RANK SUM TESTS, we determined that the Rhine sculpins didn't differ among the sampling sites and subsequently tested for host genetic differences between rhen and *C. perifretum* (peri). In

the second experiment the data didn't meet the normality requirements, even after log transformation, so we applied non-parametric KRUSKAL-WALLIS RANK SUM TESTS with genotype and population as relevant factors. All statistical analyses were carried out in the R environment (R Core Team 2015).

Results

Parasite abundance in wild caught sculpins varies with host genotype and is positively correlated with body size

In total, 27 fish, belonging to two distinct sculpins lineages, were caught at three distinct sampling sites along the contact zone Sieg/Broel, near Bonn, Germany. By examining them for parasite infection, we recorded 11 different parasite taxa (Table 4.2). The populations of the Rhine sculpins from Wahnbach upstream and Broel didn't significantly vary in parasite abundance (WILCOXON RANK SUM TEST $W = 49.5$, $P \geq 0.05$). Accordingly, we only included the factors fish genotype: *C. rhenanus* (rhen) and invasive *C.* (inv) and body size in the linear model which resulted in significant effects of fish genotype and body size, the larger the fish the higher the abundance (Figure 4.3, Table 4.3). *Gyrodactylus* sp., *Contracaecum* sp. and *Pomphorhynchus laevis* were missing in stream Broel, whereas *Diplostomum* sp. and *Crepidostomum* sp. were exclusively detected in fish from Broel.

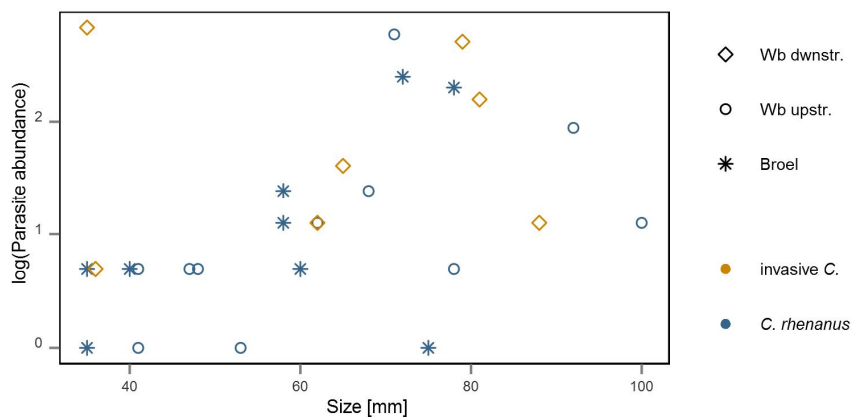


Figure 4.3 Parasite abundance of wild caught sculpins by size. The parasite abundance was shaped by size and genotype (invasive *C.* = yellow, *C. rhenanus* = blue). There was no difference between the sampling sites Wahnbach downstream (diamond), Wahnbach upstream (circle) and Broel (asterisk).

In Wahnbach upstream we didn't obtain *Contracaecum* sp, *Raphidascaris* sp. and *Cyatocephalus truncates*. Fish from Wahnbach downstream were lacking Trichodina. The highest overall abundance had the trematode *Apatemon* spp.: we determined 51 metacercariae, in total. Most of them were found in invasive sculpins from Wahnbach downstream and 16 in one individual fish. The second most abundant parasite was the acanthocephalan worm *P. laevis*: we identified 30 individuals, all in fish from Wahnbach, of which 15 were

found in one individual Rhine sculpin from Wahnbach upstream. Another major parasite was the nematode *Raphidascaris* sp. (Table 4.2).

Table 4.2: Prevalence and mean abundance of Parasites in Rhine and invasive sculpins caught at three distinct sampling sites.

	<i>C. rhenanus</i>				invasive <i>C.</i>	
	Broel n = 9		Wb upstream n = 11		Wb downstream n = 7	
	P [%]	M	P [%]	M	P [%]	M
Ciliata						
Ciliata unclassified	22	ND	18	ND	14	ND
<i>Trichodina</i>	33	ND	36	ND	0	ND
<i>Apiosoma</i>	56	ND	45	ND	14	ND
Trematoda						
<i>Gyrodactylus</i> sp	0	0.00	18	0.18	29	0.29
<i>Apatemon</i> sp MC i.o.	11	0.33	0	0.00	14	0.57
<i>Apatemon</i> sp MC b.w.	11	0.78	9	0.09	29	2.14
<i>Apatemon</i> MC	0	0.00	0	0.00	71	3.00
<i>Diplostomum</i> sp	11	0.11	0	0.00	0	0.00
<i>Crepidostomum</i> sp	11	0.22	0	0.00	0	0.00
Cestoda						
<i>Cyathocephalus truncatus</i>	22	0.22	0	0.00	14	0.14
Secernentea						
<i>Raphidascaris</i> sp	44	1.33	0	0.00	29	0.29
<i>Contraecaecum</i> sp	0	0.00	0	0.00	14	0.14
Palaeacanthocephala						
<i>Pomphorhynchus laevis</i>	0	0.00	64	2.64	14	0.14

P = prevalence; M = mean; ND = not determined; MC = metacercariae; i.o. = inner organ; b.w. = body wall; Wb = Wahnbach

Table 4.3: Analysis of variance of parasite abundance of wild caught Rhine and invasive sculpins. The full model contained the term genotype (rhen, inv), sex (f, m) and size. The parasite abundance was significantly influenced by host genotype and size but not by sex.

Term	Df	Sum Sq	Mean Sq	F value	Pr (>F)
genotype	1	2.773	2.7735	4.686	0.0406
size	1	2.741	2.7406	4.630	0.04017
Residuals	24	14.205	0.5919		

Bold P-values indicate significance ($P < 0.05$).

Infection experiment revealed that susceptibility to *D. pseudospathaceum* in part differs between distinct host populations

Since it was not possible to disentangle environmental from host genetic effects in the previous study we performed two infection experiments in the laboratory. The goal was to investigate whether sculpins from different families and genetic lineages vary in susceptibility to the common parasite *D. pseudospathaceum*. For this purpose, fish were exposed to 100 cercariae of *D. pseudospathaceum*, each, and the infection rate was assessed after 24 h by dissecting the eye lenses of the fish.

In the first experiment, 14 lab bred sculpins, originating from two different populations of

C. rhenanus (Broel n = 5, Naaf n = 3), from one population of *C. perifretum* (Laarse Beek n = 4) and two F1 crosses (*C. rhenanus* from Wahnbach x *C. perifretum* from Laarese Beek), were infected with cercariae of *D. pseudopathaceum* which revealed no significant differences between Rhine sculpins from Broel and Naaf (WILCOXON RANK SUM TEST, $W = 5$, $P = 0.549$) and further not between all Rhine and Scheldt sculpins (WILCOXON RANK SUM TEST, $W = 18$, $P = 0.799$) (Figure 4.4). The overall parasite abundance varied from 2 to 28 cercariae per fish.

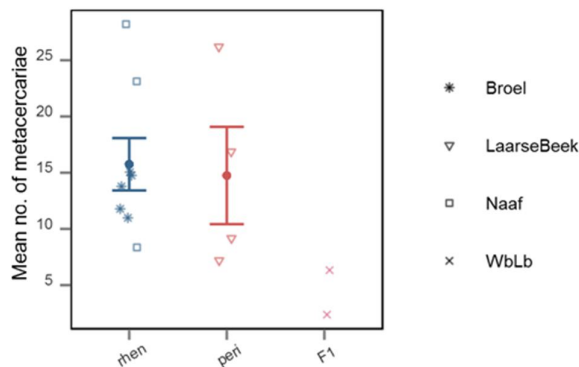


Figure 4.4: Infection rate of *D. pseudopathaceum* in lab bred sculpins. Rhine sculpins originating from stream Broel (asterisk, n = 5) or Naaf (square, n = 3), Scheldt sculpins from stream Laarse Beek (triangle, n = 4) and 2 F1 cross (cross) were exposed to 100 cercariae, each, for 24 h. The F1 were excluded from statistical analyses. The overall parasite abundance varied but there were no significant differences between population origins and host genotypes ($P \geq 0.05$).

In the second experiment 52 lab bred fish, two families of invasive *C.* (inv14 n = 12, inv21 n = 12) and three families of *C. perifretum* (LB2 n = 8, LB3 n = 8, WN n = 12), were exposed to 100 cercariae of *D. pseudopathaceum* which resulted in significantly different parasite abundance depending on family origin (KRUSKAL-WALLIS RANK SUM TEST, $X^2 = 26.071$, $P < 0.001$). The infection rate of family inv21 was decreased, whereas fish from WN were most infected (Figure 4.5). Generally, the fish varied greatly in their parasite abundance. We counted 7 to 59 individuals per fish and the highest infection success in Scheldt sculpins from Witte Nete (WN) and the lowest in invasive sculpins from Pleisbach (inv21) (Figure 4.5). There was no significant difference between fish genotype (KRUSKAL-WALLIS RANK SUM TEST, $X^2 = 1.3605$, $P = 0.244$).

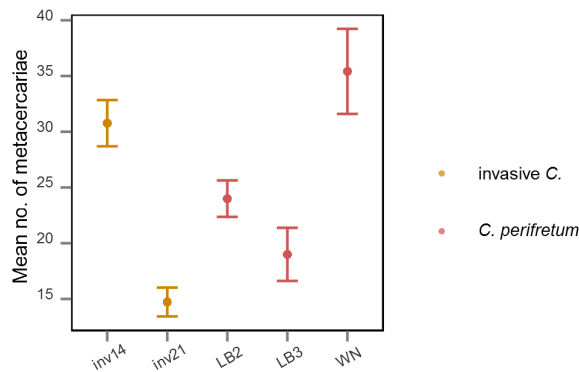


Figure 4.5: Infection rate of *D. pseudospathaceum* in lab bred sculpins. Two families of invasive *C.* inv14 ($n = 12$) and inv21 ($n = 12$) and three families of *C. perifretum*, LB2 ($n = 8$), LB3 ($n = 8$) and WN ($n = 12$) were exposed to 100 cercariae for 24 h, each. The parasite abundance differed significantly among the population origin ($P < 0.001$) but there was no host lineage specific effect ($P \geq 0.5$).

Discussion

There's recent interest in ecological speciation and the emergence of reproductive barriers (Schluter 2001, 2009; Rundle & Nosil 2005; Hendry *et al.* 2007; Nosil 2012) but parasite-mediated divergent selection has gained little attention in this context (Karvonen & Seehausen 2012). Parasites can have great influence on their hosts' behavior and immune system (Price *et al.* 1986; Wood *et al.* 2007) and thus variation in parasite infection rates between host populations can emerge through genetic differentiation. Moreover, a parasite's abundance depends on environmental factors in the respective habitat, like temperature, pH and intermediate and final host community structure (Arneberg *et al.* 1998; Karvonen *et al.* 2010). Hence, it can be expected that different streams and rivers and the fish therein, although not very distant from each other, differ in parasite richness and diversity.

The present pilot studies aimed at investigating first the parasite composition of two different sculpin lineages inhabiting secondary contact zones near Bonn, Germany, and second the susceptibility of distinct sculpin families, to the common fish parasite *D. pseudospathaceum*. Differences between host populations could indicate a potential selective impact of the parasites on the host which may limit dispersal to foreign habitats. In wild sculpins, we identified three ciliate species and eight metazoan parasite taxa. There was a significant effect of host genotype and size on parasite abundance (Table 4.3) but no significant difference between the two sampling sites at the Rhine. In the first infection experiment I didn't detect any significant difference between Scheldt and Rhine sculpins and none among the families. However, in the second infection experiment the parasite susceptibility significantly varied between the five host families but not between the genotypes invasive and Scheldt sculpin.

Parasites in wild Cottus

The streams of the Sieg system, although very close to each other, may differ by intermediate and final hosts and various environmental factors the parasite is exposed to. These fluctuating patterns can influence the parasite's physiology and distribution (Kalbe *et al.* 2002; Vale *et al.* 2008). In this study we didn't measure any biotic factor, like salinity, temperature or pH and thus can only speculate that the streams differ and thereby potentially influence the parasite communities and transmission in them. This assumption is supported by variation between the respective parasite species among the sampling sites and host genotypes. In fact, we cannot disentangle these two factors with the present study, because we obtained different host genotypes inhabiting distinct streams. The results don't permit to conclude whether occurring variation in parasite abundance and diversity were caused by genotype or habitat effects. Apart from the trematode *Apatemon* sp. and two ciliates all parasites only occur at one or two sites. Fish from Wahnbach upstream contained remarkably few parasites (Table 4.2). However, the parasite diversity between the sampling sites was not significantly different (data not shown). We determined that larger fish contained more parasites but since size is, to a certain extent, correlated with age we conclude that the larger fish were on average older than the smaller ones. It has been shown before that parasites accumulate in fish over time and that older fish have generally more parasites than younger fish because they eat more infected prey and further provide more space for the parasite (Poulin 2007). This could be related to our system, either. The fish examined here had generally lower parasite abundance than fish in other studies (Kalbe *et al.* 2002; Wegner *et al.* 2003; MacColl & Chapman 2010; Karvonen *et al.* 2013a, 2015) which could have been caused by seasonal impact. The sculpins have been caught in late February and the water temperature was approximately 4-8°C as determined by the 'Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen' under <http://luadb.lids.nrw.de /LUA/hygon/pegel.php> (Station 'Broel' N 50°47'9.8 E 7°19'11.8; accessed 03/08/16). Several studies assessed that temperature has a great impact on development, physiology and behavior of parasites and accordingly can lead to seasonality of parasite infections in general and particularly in fish (Altizer *et al.* 2006; Karvonen *et al.* 2010; Antonelli *et al.* 2015): the warmer the water the higher the parasite load. Among the parasites being identified the trematode *Apatemon* sp. had the highest abundance and was mostly found in invasive sculpins. Similar to *D. pseudospathaceum*, cercariae of *Apatemon* sp develop in snails, enter the fish through the skin, become metacercariae and 'wait' that the fish gets eaten by their final host – a fish eating bird (Vojtek 1964; Blair 1976). At high infection intensities the parasite can cause mortality of the fish host (Gordon & Rau 1982; Ieshko *et al.* 2013). The second most abundant parasite in the sculpins was *P. laevis* which is one of the most frequent and widely distributed acanthocephalan fish parasite in Europe and is well known to infect *Cottus* sp. (Rumpus 1975; Lagrue *et al.* 2007). Although many individuals of *P. laevis* can accumulate in one single fish they cause local damage by penetration but no severe general effects (Hine & Kennedy 1974). Taken together, we cannot answer whether parasite-mediated selection is shaping the contact zone with the present data of wild caught fish. A higher sample size in the particular fish

genotypes would be required as well as sampling along the contact zone, as it has been done in Chapter I.

Parasite infections in lab-raised Cottus spp.

In addition to the examination of parasite communities in wild caught sculpins we performed two infection experiments wherein we exposed distinct sculpin genotypes to an equal amount of cercariae of the common fish parasite *D. pseudospathaceum*, each, and assessed their infection rate after 24 h. The setup mimics a possible scenario in nature, the amount of cercariae is reasonable. *D. pseudospathaceum* can affect the host's vision and thereby increase the probability for the fish to be eaten by a bird, the final host of *D. pseudospathaceum*. Thus, the selection pressure on the host to develop avoidance behavior or enhance its immune defense is high. However, we didn't detect significant differences in the first experiment, neither between the habitats the Rhine sculpins originate from nor between the fish genotypes Rhine and Scheldt sculpins. One possible reason for this could be the lack of statistical power since. In the second infection experiment we compared two families of invasive with three families of Scheldt sculpins and determined that family is mainly influencing the infection rate but not fish genotype. The Scheldt sculpin families from stream Laarse Beek had very similar and both medium parasite abundance, whereas the fish from stream Witte Nete had the highest infection rate. The two invasive families differed in their abundance and had the lowest and the second highest parasite infection rate (Figure 4.5). However, the sample size in the respective groups was rather low and additional in-depth studies would be needed. It has been determined that sticklebacks, which are regularly exposed to *D. pseudospathaceum* in their natural habitat, were less susceptible than those from habitats where the parasite is absent (Kalbe & Kurtz 2006). Further, the authors identified significant difference in immunocompetence and energy status between the fish groups. The less susceptible sticklebacks were in better shape and had a higher respiratory burst activity. The intermediate host of *D. pseudospathaceum*, the snail *L. stagnalis*, is probably not very frequent in rivers, as it is mostly distributed in lakes. Thus, one would expect high infection rates in sculpins, similar to sticklebacks from rivers, because they are not frequently exposed to this parasite. However, the opposite was true for our system. There's evidence that *Diplostomum* sp. infects *Cottus* sp. (Heckmann 1983; Muzzall & Bowen 2002; Ieshko *et al.* 2013) but there's, to our knowledge, no data for *C. rhenanus* or *C. perifretum*.

Conclusions

Generally, these pilot studies served to get a first insight into parasite communities in Western European sculpins and additionally to identify whether significant variation in parasite abundance and diversity suggest promising results for larger in-depth studies. In sum, all sculpins had low parasite abundance in the wild and further low infection rates in the lab. However, the sample sizes of the respective groups were low and in-depth sampling of several families per fish genotype and location would increase the statistical power and presumably differences could be identified. Simultaneously, we examined the bacterial

communities of distinct sculpin lineages and determined significant effects with fewer samples than we would have needed to identify significant differences in parasite communities. Hence, it was, in the end, more promising for us to continue studying bacterial microbiota as drivers to adaptive speciation in the *Cottus* system.

Supplement

Tables

Table 4.S1: Measured variable of wild caught Rhine and invasive sculpins from Broel, Wahnbach upstream (Wbu) and downstream (Wbd). Fish and organs were weighted to the nearest 0.1 mg and length was measured to the nearest mm.

Origin	Genotype	Sex	Weigh [g]t	SL [mm]	TL [mm]	Liver [g]	Spleen [g]	Parasite abundance
Broel	rhen	m	0.8409	49	40	0.0098	0.0001	1
Broel	rhen	f	3.6218	72	60	0.1322	0.0031	1
Broel	rhen	f	3.4669	71	58	0.0717	0.005	2
Broel	rhen	m	0.6531	43	35	0.078	0.0003	0
Broel	rhen	f	0.5823	43	35	0.0113	0.0003	1
Broel	rhen	f	3.7028	69	58	0.1226	0.0016	3
Broel	rhen	m	7.6869	90	75	0.1093	0.0032	0
Broel	rhen	m	10.1366	96	78	0.1192	0.0043	9
Broel	rhen	f	7.5179	85	72	0.2093	0.0014	10
Wbu	rhen	f	7.5925	92	78	0.2565	0.004	1
Wbu	rhen	m	2.0845	58	48	0.0314	0.0092	1
Wbu	rhen	f	4.3235	74	62	0.1396	0.0079	2
Wbu	rhen	f	1.0115	50	41	0.0213	0.0011	0
Wbu	rhen	m	2.5834	64	53	0.0278	0.0018	0
Wbu	rhen	m	0.9665	50	41	0.0139	0.0003	1
Wbu	rhen	m	15.2222	105	92	0.1672	0.0143	6
Wbu	rhen	m	1.5285	57	47	0.0221	0.0012	1
Wbu	rhen	f	6.4279	84	71	0.1557	0.0025	15
Wbu	rhen	f	6.2887	83	68	0.16	0.0045	3
Wbu	rhen	m	15.6918	112	100	0.1387	0.0216	2
Wbd	inv	f	0.9307	45	36	0.0327	0.001	1
Wbd	inv	m	6.068	77	65	0.1287	0.0082	4
Wbd	inv	f	0.7552	43	35	0.057	0.0004	16
Wbd	inv	f	10.257	96	81	0.243	0.0084	8
Wbd	inv	f	14.2481	95	79	0.3826	0.0059	14
Wbd	inv	m	4.7667	73	62	0.057	0.0112	2
Wbd	inv	f	12.9445	98	88	0.3193	0.0084	2

SL = Standard Length; TL = Total Length

Table 4.S2: Analysis of variance of infection rate within *C. rhenanus* in experiment I. The full model contained the terms population origin (Broel, Naaf) and weight. None of the factors explained variation in the data.

Term	Df	Sum Sq	Mean Sq	F value	Pr (>F)
origin	1	73.633	73.633	1.6242	0.2585
Residuals	1	3.188	3.188	0.0703	0.8014
Residuals	1	226.678	45.336		

Table 4.S3: Analysis of variance of parasite abundance of wild caught Rhine and Scheldt sculpins. The full model contained the term genotype (*rhen*, *perl*). There was no significant difference between the fish genotypes.

Term	Df	Sum Sq	Mean Sq	F value	Pr (>F)
genotype	1	0.03444	0.034443	0.1815	0.6791
Residuals	10	1.89775	0.189775		

CONCLUSIONS AND PERSPECTIVES

It is evident that host-bacterial interactions are a complex system (defined as holobiont (Margulis 1993)) shaped by gene-by-gene-by-environmental interactions. However, the role of host-associated bacteria and parasites on the emergence of reproductive isolation to the point of speciation has gained little attention (Karvonen & Seehausen 2012; Brucker & Bordenstein 2012) and is largely unknown. In this thesis I examined if bacterial microbiota and parasite communities are shaped by host genetic factors, indicating that they may be drivers of adaptive speciation processes in a secondary contact zone of European freshwater *Cottus*. In the following I will relate the chapters to each other and will suggest further perspectives which would extend the understanding of the role of bacteria and parasites on speciation processes.

Particular focus on host genetic effects

By characterizing naturally hybridizing *C. rhenanus* and invasive *Cottus*, I determined a main impact of host habitat and a minor influence of host genotype on the bacterial community structure. In contrast, I identified a major role of host genotype and secondary impact of host habitat when analyzing *B. barbatula* and *P. phoxinus*. However, the putative discrepancy between Chapter I and II only reflects the complexity and variability of host-associated bacterial microbiota in general and in particular in the present system. To be able to investigate the respective influence of host genetic and environment/habitat it is necessary to disentangle the factors from each other. With laboratory experiments it is possible to control for environmental variation and only consider genetic differences. But there's the risk that by chance one host genotype is better adapted to laboratory environments which could bias the results. Moreover, laboratory conditions are artificial and thus emerging variations don't reflect the actual circumstances in nature.

Naturally emerged hybrid zones, wherein admixed and different pure host genotypes occur in two distinct habitats and under natural selection pressures seem to be the perfect system to disentangle genetic from environmental factors. In the present study, I indeed examined *C. rhenanus* and invasive *C.* occurring in both habitats stream and river, each. However, the sample size of the respective fish in foreign habitats was low and their dispersal was limited to approximately 2 km distant the contact zone, each. Hence, genetic factors were not completely independent from habitat effects. The minor influence of host genotype on bacterial community variation I have identified in chapter I, is presumably a result of this limitation. This gave rise to the second study (Chapter II), investigating *B. barbatula* and *P. phoxinus* occurring in both habitats stream and river in high abundance. Results indicate a major host genetic effect on skin bacteria abundance and diversity: differences between distinct fish species in the same habitat were larger than between the same host species in distinct habitats. Admittedly, this does not reveal the role of microbiota for the *Cottus* system but it does show, that the choice of fin tissue in this study represents a useful approach to detect associations of bacteria with host genotypes. Hence, by using

Conclusions and perspectives

larger sample size of *Cottus* at all sampling sites, as well, future work would enable to examine more pure host genotypes of *Cottus* in foreign habitats. Increased statistical power should then reveal host genetic effects on bacterial community variation, like I obtained them for *B. barbatula* and *P. phoxinus*.

It has been demonstrated that QTL analysis is a useful tool to identify candidate loci underlying certain traits or phenotypes in the host, like for instance bacterial composition or even the presence/absence of single bacteria (Benson *et al.* 2010; Srinivas *et al.* 2013; Boutin *et al.* 2014; Wang *et al.* 2015). Usually, QTL studies use laboratory-generated F2 crosses for mapping but their genetic and phenotypic variation is low in comparison to natural variation. Instead of artificial crosses naturally emerged recombinants generated by hybridization between divergent parental taxa have been used for genetic mapping for several decades (Hewitt 1988; Rieseberg & Buerkle 2002; Buerkle & Lexer 2008). These recombinants have been created by many generations of intercrossing and therefore the mapping resolution is supposed to be high. Future work could use the *Cottus* hybrid zone for QTL mapping to detect candidate loci or regions being associated with bacterial community structure as it has been done for other traits (e.g. Turner & Harr 2014; Pallares *et al.* 2015).

However, QTL that are not selected for in a natural hybrid zone and contribute to isolation occur at low frequency and thus are difficult to detect (Rieseberg & Buerkle 2002). Moreover, using wild individuals impedes to control for all possible variances influencing bacterial abundance and diversity. Hence, it might further be useful to perform mapping studies with laboratory-bred intercrosses, as it has been demonstrated in mice (Benson *et al.* 2010; Srinivas *et al.* 2013; Wang *et al.* 2015). Future work could include several advanced generations of large mapping families. This would provide knowledge about how host genetic variation contributes to variation in bacterial communities.

Environmental impact on bacterial communities

Besides focusing on genetic factors examining influential ecological effects causing variation in bacterial communities would help understanding the complex host-symbiont interactions in *Cottus*. My results have shown that the bacterial composition is habitat-dependent but apart from constant temperature records by the 'Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen', accessible under <http://luadb.lids.nrw.de/LUA/hygon/pegel.php>, no further information about ecological differences between stream and river is available. Thus, considering further measurements in the respective habitats is needed to better understand the factors underlying host-bacterial community dynamics. Besides temperature (Erwin *et al.* 2012; Lokmer & Wegner 2015; Neuman *et al.* 2016), particularly pH and salinity are known to influence skin bacterial and parasite abundance (Fierer & Jackson 2006; Lozupone & Knight 2007; Roeselers *et al.* 2011; Sullam *et al.* 2012). Also the food resources differ between the habitats which could be related to variation in bacterial composition (Sullam *et al.* 2012).

Relation of host-pathogen interactions to adaptive immune response

Contrasting environments likely harbor distinct parasites and pathogens which requires contrasting adaptive immune responses in the hosts. A reciprocal transplant experiment, wherein individuals of both populations/genotypes are exposed to either environment, would enable to analyze local adaptation for instance in the MHC immune genes. It has been suggested that parasites in contrasting environments drive divergent selection on locally adapted MHC alleles (Eizaguirre & Lenz 2010; Eizaguirre *et al.* 2012a) and influence MHC composition e.g. in fish (Reusch *et al.* 2001; e.g. Wegner *et al.* 2003; Eizaguirre *et al.* 2012b). *C. rhenanus*, invasive *C.* and hybrids could be placed in cages into the stream and into the river, distant from the contact zone. After several months, the fish would be sampled and their bacteria and parasite communities analyzed. Results would identify whether varying bacteria/parasite composition stem from host genetic differences or from the respective habitat fauna. Occurring variation could then be related to MHC allele composition.

Conclusion

In sum, this thesis confirms the importance of considering ecological and evolutionary processes as a whole complex consisting of numerous components all interacting with each other. I cannot say whether bacteria or parasites are driving early adaptive speciation processes in *Cottus* but I demonstrated that they vary between habitats and genotypes, which is in agreement with ideas suggesting a considerable impact on their hosts. Thus, I can also not exclude that these symbionts play a role in limiting dispersal to foreign habitats. More work is needed to better understand how species evolve and which factors are underlying these processes.

REFERENCES

- Abbott R, Albach D, Ansell S *et al.* (2013) Hybridization and speciation. *Journal of Evolutionary Biology*, 26, 229–246.
- Alegado RA, King N (2014) Bacterial Influences on Animal Origins. *Cold Spring Harbor Perspectives in Biology*, 6, a016162.
- Alexander AD, Orcutt RP, Henry JC *et al.* (2006) Quantitative PCR assays for mouse enteric flora reveal strain-dependent differences in composition that are influenced by the microenvironment. *Mammalian Genome*, 17, 1093–1104.
- Altizer S, Dobson A, Hosseini P *et al.* (2006) Seasonality and the dynamics of infectious diseases. *Ecology Letters*, 9, 467–484.
- Anderson KE, Russell JA, Moreau CS *et al.* (2012) Highly similar microbial communities are shared among related and trophically similar ant species. *Molecular Ecology*, 21, 2282–2296.
- Antonelli L, Foata J, Quilichini Y, Marchand B (2015) Influence of season and site location on European cultured sea bass parasites in Corsican fish farms using indicator species analysis (IndVal). *Parasitology Research*, 115, 561–568.
- Arneberg P, Skorping A, Grenfell B, Read AF (1998) Host densities as determinants of abundance in parasite communities. *Proceedings of the Royal Society of London B: Biological Sciences*, 265, 1283–1289.
- Arnold ML (1992) Natural Hybridization as an Evolutionary Process. *Annual Review of Ecology and Systematics*, 23, 237–261.
- Austin B (2006) The Bacterial Microflora of Fish, Revised. *The Scientific World Journal*, 6, 931–945.
- Baird SJE, Ribas A, Macholan M *et al.* (2012) Where Are the Wormy Mice? A Reexamination of Hybrid Parasitism in the European House Mouse Hybrid Zone. *Evolution*, 66, 2757–2772.

References

- Bakke TA, Jansen PA, Hansen LP (1990) Differences in the host resistance of Atlantic salmon, *Salmo salar* L., stocks to the monogenean *Gyrodactylus salaris* Malmberg, 1957. *Journal of Fish Biology*, 37, 577–587.
- Balcazar JL, de Blas I, Ruiz-Zarzuola I *et al.* (2006) The role of probiotics in aquaculture. *Veterinary Microbiology*, 114, 173–186.
- Barton NH, Hewitt GM (1985) Analysis of Hybrid Zones. *Annual Review of Ecology and Systematics*, 16, 113–148.
- Benhamed S, Guardiola FA, Mars M, Esteban MÁ (2014) Pathogen bacteria adhesion to skin mucus of fishes. *Veterinary Microbiology*, 171, 1–12.
- Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289–300.
- Benson AK, Kelly SA, Legge R *et al.* (2010) Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proceedings of the National Academy of Sciences*, 107, 18933–18938.
- Bergmann GT, Bates ST, Eilers KG *et al.* (2011) The under-recognized dominance of Verrucomicrobia in soil bacterial communities. *Soil Biology and Biochemistry*, 43, 1450–1455.
- Bernadsky G, Sar N, Rosenberg E (1993) Drag reduction of fish skin mucus: Relationship to mode of swimming and size. *Journal of Fish Biology*, 42, 797–800.
- Bevins CL, Salzman NH (2011) The potter's wheel: the host's role in sculpting its microbiota. *Cellular and Molecular Life Sciences*, 68, 3675–3685.
- Blair D (1976) Observations on the life-cycle of the strigeoid trematode, *Apatemon (Apatemon) gracilis* (Rudolphi, 1819) Szidat, 1928. *Journal of Helminthology*, 50, 125–132.

- Bolnick DI, Snowberg LK, Patenia C *et al.* (2009) Phenotype-Dependent Native Habitat Preference Facilitates Divergence between Parapatric Lake and Stream Stickleback. *Evolution*, 63, 2004–2016.
- Bordenstein SR, Theis KR (2015) Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes. *PLoS Biol*, 13, e1002226.
- Boutin S, Bernatchez L, Audet C, Derome N (2013) Network Analysis Highlights Complex Interactions between Pathogen, Host and Commensal Microbiota. *Plos One*, 8.
- Boutin S, Sauvage C, Bernatchez L, Audet C, Derome N (2014) Inter Individual Variations of the Fish Skin Microbiota: Host Genetics Basis of Mutualism? *PLoS ONE*, 9, e102649.
- Brucker RM, Bordenstein SR (2012) Speciation by symbiosis. *Trends in Ecology & Evolution*, 27, 443–451.
- Brucker RM, Bordenstein SR (2013) The Hologenomic Basis of Speciation: Gut Bacteria Cause Hybrid Lethality in the Genus *Nasonia*. *Science*, 341, 667–669.
- Buerkle CA, Lexer C (2008) Admixture as the basis for genetic mapping. *Trends in Ecology & Evolution*, 23, 686–694.
- Buerkle CA, Morris RJ, Asmussen MA, Rieseberg LH (2000) The likelihood of homoploid hybrid speciation. *Heredity*, 84 (Pt 4), 441–451.
- Burke JM, Arnold ML (2001) Genetics and the Fitness of Hybrids. *Annual Review of Genetics*, 35, 31–52.
- Caceres MD, Legendre P (2009) Associations between species and groups of sites: indices and statistical inference. *Ecology*, 90, 3566–3574.
- Cahill MM (1990) Bacterial Flora of Fishes: A Review. *Microbial Ecology*, 19, 21–41.
- Campbell JH, Foster CM, Vishnivetskaya T *et al.* (2012) Host genetic and environmental effects on mouse intestinal microbiota. *The ISME Journal*, 6, 2033–2044.

References

- Caporaso JG, Lauber CL, Walters WA *et al.* (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, 108, 4516–4522.
- Carius HJ, Little TJ, Ebert D (2001) Genetic Variation in a Host-Parasite Association: Potential for Coevolution and Frequency-Dependent Selection. *Evolution*, 55, 1136–1145.
- Chaucheyras-Durand F, Durand H (2009) Probiotics in animal nutrition and health. *Beneficial Microbes*, 1, 3–9.
- Cheng J, Czypionka T, Nolte AW (2013) The genomics of incompatibility factors and sex determination in hybridizing species of *Cottus* (Pisces). *Heredity*.
- Cheng J, Sedlazeck F, Altmüller J, Nolte AW (2015) Ectodysplasin signalling genes and phenotypic evolution in sculpins (*Cottus*). *Proc. R. Soc. B*, 282, 20150746.
- Chiarello M, Villéger S, Bouvier C, Bettarel Y, Bouvier T (2015) High diversity of skin-associated bacterial communities of marine fishes is promoted by their high variability among body parts, individuals and species. *FEMS Microbiology Ecology*, 91, fiv061.
- Chubb JC, Seppälä T, Lüscher A, Milinski M, Valtonen ET (2006) *Schistocephalus cotti* n. sp. (Cestoda: Pseudophyllidea) plerocercoids from bullheads *Cottus gobio* L. in an Arctic river in Finland, with a key to the plerocercoids of the Palaearctic species of the genus. *Systematic Parasitology*, 65, 161–170.
- Cogen A I., Nizet V, Gallo R I. (2008) Skin microbiota: a source of disease or defence? *British Journal of Dermatology*, 158, 442–455.
- Costello EK, Lauber CL, Hamady M *et al.* (2009) Bacterial Community Variation in Human Body Habitats Across Space and Time. *Science (New York, N.Y.)*, 326, 1694–1697.
- Coyne JA, Orr HA (2004) Speciation.

- Crowden AE, Broom DM (1980) Effects of the eye-fluke, *Diplostomum spathaceum*, on the behaviour of dace (*Leuciscus leuciscus*). *Animal Behaviour*, 28, 287–294.
- Crump BC, Hobbie JE (2005) Synchrony and seasonality in bacterioplankton communities of two temperate rivers. *Limnology and Oceanography*, 50, 1718–1729.
- Cryan JF, Dinan TG (2012) Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nature Reviews Neuroscience*, 13, 701–712.
- Czypionka T, Cheng J, Pozhitkov A, Nolte AW (2012) Transcriptome changes after genome-wide admixture in invasive sculpins (*Cottus*). *Molecular Ecology*, 21, 4797–4810.
- Czypionka T, Krugman T, Altmüller J *et al.* (2015) Ecological transcriptomics – a non-lethal sampling approach for endangered fire salamanders. *Methods in Ecology and Evolution*, 6, 1417–1425.
- Darwin C (1859) *On the Origin of Species by Means of Natural Selection or the Preservation of Favored Races in the Struggle for Life*. J. Murray, London.
- David LA, Maurice CF, Carmody RN *et al.* (2014) Diet rapidly and reproducibly alters the human gut microbiome. *Nature*, 505, 559–563.
- De Caceres M, Legendre P, Wiser SK, Brotons L (2012) Using species combinations in indicator value analyses. *Methods in Ecology and Evolution*, 3, 973–982.
- De Nie HW de (1997) Atlas van de Nederlandse zoetwatervissen.
- De Queiroz K (2007) Species Concepts and Species Delimitation. *Systematic Biology*, 56, 879–886.
- Dobzhansky T (1937) *Genetics and the Origin of Species*. Columbia University Press.
- Dominguez-Bello MG, Costello EK, Contreras M *et al.* (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, 11971–11975.

References

- Dufrene M, Legendre P (1997) Species Assemblages and Indicator Species: The Need for a Flexible Asymmetrical Approach. *Ecological Monographs*, 67, 345–366.
- Duncan SH, Louis P, Thomson JM, Flint HJ (2009) The role of pH in determining the species composition of the human colonic microbiota. *Environmental Microbiology*, 11, 2112–2122.
- Ebert D (2008) Host–parasite coevolution: Insights from the *Daphnia*–parasite model system. *Current Opinion in Microbiology*, 11, 290–301.
- Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, 27, 2194–2200.
- Eizaguirre C, Lenz TL (2010) Major histocompatibility complex polymorphism: dynamics and consequences of parasite-mediated local adaptation in fishes. *Journal of Fish Biology*, 77, 2023–2047.
- Eizaguirre C, Lenz TL, Kalbe M, Milinski M (2012a) Rapid and adaptive evolution of MHC genes under parasite selection in experimental vertebrate populations. *Nature Communications*, 3, 621.
- Eizaguirre C, Lenz TL, Kalbe M, Milinski M (2012b) Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. *Ecology Letters*, 15, 723–731.
- Eizaguirre C, Yeates SE, Lenz TL, Kalbe M, Milinski M (2009) MHC-based mate choice combines good genes and maintenance of MHC polymorphism. *Molecular Ecology*, 18, 3316–3329.
- Engelstädter J, Hurst GDD (2009) The Ecology and Evolution of Microbes that Manipulate Host Reproduction. *Annual Review of Ecology, Evolution, and Systematics*, 40, 127–149.

- Englbrecht CC, Freyhof J, Nolte A *et al.* (2000) Phylogeography of the bullhead *Cottus gobio* (Pisces: Teleostei: Cottidae) suggests a pre-pleistocene origin of the major central European populations. *Molecular Ecology*, 9, 709–722.
- Erwin PM, Pita L, López-Legentil S, Turon X (2012) Stability of Sponge-Associated Bacteria over Large Seasonal Shifts in Temperature and Irradiance. *Applied and Environmental Microbiology*, 78, 7358–7368.
- Esteban MÁ (2012) An Overview of the Immunological Defenses in Fish Skin. *ISRN Immunology*, 1–29.
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multi-locus genotype data: linked loci and correlated allele frequencies. *Genetics*, 164, 1567–1587.
- Fierer N, Jackson RB (2006) The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 626–631.
- Fitzpatrick BM, Allison AL (2014) Similarity and differentiation between bacteria associated with skin of salamanders (*Plethodon jordani*) and free-living assemblages. *FEMS Microbiology Ecology*, 88, 482–494.
- Fjellheim AJ, Playfoot KJ, Skjermo J, Vadstein O (2012) Inter-individual variation in the dominant intestinal microbiota of reared Atlantic cod (*Gadus morhua* L.) larvae. *Aquaculture Research*, 43, 1499–1508.
- Franzenburg S, Walter J, Kuenzel S *et al.* (2013) Distinct antimicrobial peptide expression determines host species-specific bacterial associations. *Proceedings of the National Academy of Sciences of the United States of America*, 110, E3730–E3738.
- Fraser BA, Neff BD (2009) Parasite mediated homogenizing selection at the MHC in guppies. *Genetica*, 138, 273–278.

References

- Fraune S, Bosch TCG (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, 13146–13151.
- Fraune S, Bosch TCG (2010) Why bacteria matter in animal development and evolution. *BioEssays*, 32, 571–580.
- Freyhof J, Kottelat M, Nolte A (2005) Taxonomic diversity of European *Cottus* with description of eight new species (Teleostei: Cottidae).
- Galindo-Villegas J, García-Moreno D, Oliveira S de, Meseguer J, Mulero V (2012) Regulation of immunity and disease resistance by commensal microbes and chromatin modifications during zebrafish development. *Proceedings of the National Academy of Sciences*, 109, E2605–E2614.
- Givens CE, Ransom B, Bano N, Hollibaugh JT (2015) Comparison of the gut microbiomes of 12 bony fish and 3 shark species. *Marine Ecology Progress Series*, 518, 209–223.
- Gomez D, Sunyer JO, Salinas I (2013) The mucosal immune system of fish: The evolution of tolerating commensals while fighting pathogens. *Fish & Shellfish Immunology*, 35, 1729–1739.
- Gompert Z, Buerkle CA (2009) A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. *Molecular Ecology*, 18, 1207–1224.
- Goodrich JK, Waters JL, Poole AC *et al.* (2014) Human Genetics Shape the Gut Microbiome. *Cell*, 159, 789–799.
- Gordon DM, Rau ME (1982) Possible evidence for mortality induced by the parasite *Apatemon gracilis* in a population of brook sticklebacks (*Culaea inconstans*). *Parasitology*, 84, 41–47.
- Grice EA, Kong HH, Conlan S *et al.* (2009) Topographical and Temporal Diversity of the Human Skin Microbiome. *Science*, 324, 1190–1192.
- Grice EA, Segre JA (2011) The skin microbiome. *Nature reviews. Microbiology*, 9, 244–253.

- Gross BL, Rieseberg LH (2005) The Ecological Genetics of Homoploid Hybrid Speciation. *The Journal of heredity*, 96, 241–252.
- Guttel Y, Ben-Ami F (2014) The maintenance of hybrids by parasitism in a freshwater snail. *International Journal for Parasitology*, 44, 1001–1008.
- Haldane JBS (1949) Disease and Evolution. , 19, 1–11.
- Harris PD (1985) Species of *Gyrodactylus* von Nordmann, 1832 (Monogenea: Gyrodactyli-
dae) from freshwater fishes in southern England, with a description of *Gyrodactylus roгатensis* sp. nov. from the bullhead *Cottus gobio* L. *Journal of Natural History*, 19, 791–809.
- Harrison RG (Ed.) (1993) *Hybrid zones and the evolutionary process*. Oxford University Press, New York.
- Hatfield T, Schluter D (1999) Ecological Speciation in Sticklebacks: Environment-Dependent Hybrid Fitness. *Evolution*, 53, 866–873.
- Hawkes CV, Elle AEE, Whitlock EMC (2007) Are Invaders Moving Targets? The Generality and Persistence of Advantages in Size, Reproduction, and Enemy Release in Invasive Plant Species with Time since Introduction. *The American Naturalist*, 170, 832–843.
- Heckmann R (1983) Eye Fluke (*Diplostomum spathaceum*) of Fishes from the Upper Salmon River Near Obsidian, Idaho. *The Great Basin Naturalist*, 43, 675–683.
- Hendry AP, Nosil P, Rieseberg LH (2007) The speed of ecological speciation. *Functional Ecology*, 21, 455–464.
- Hewitt GM (1988) Hybrid zones-natural laboratories for evolutionary studies. *Trends in Ecology & Evolution*, 3, 158–167.
- Hildebrand F, Nguyen TLA, Brinkman B *et al.* (2013) Inflammation-associated enterotypes, host genotype, cage and inter-individual effects drive gut microbiota variation in common laboratory mice. *Genome Biology*, 14, R4.

References

- Hine PM, Kennedy CR (1974) Observations on the distribution, specificity and pathogenicity of the acanthocephalan *Pomphorhynchus laevis* (Müller). *Journal of Fish Biology*, 6, 521–535.
- Hongoh Y, Deevong P, Inoue T *et al.* (2005) Intra- and Interspecific Comparisons of Bacterial Diversity and Community Structure Support Coevolution of Gut Microbiota and Termite Host. *Applied and Environmental Microbiology*, 71, 6590–6599.
- Hooper LV, Littman DR, Macpherson AJ (2012) Interactions Between the Microbiota and the Immune System. *Science*, 336, 1268–1273.
- Ieshko EP, Shul'man BS, Lebedeva DI, Barskaya YY, Niemela E (2013) Bullhead (*Cottus gobio* L.) invasion in the Utsjoki River (Northern Finland): Parasitological aspects. *Russian Journal of Biological Invasions*, 4, 17–23.
- Janeway C (Ed.) (2005) *Immunobiology: the immune system in health and disease*. Garland Science, New York.
- Jensen S, Ovreas L, Bergh O, Torsvik V (2004) Phylogenetic analysis of bacterial communities associated with larvae of the Atlantic halibut propose succession from a uniform normal flora. *Systematic and Applied Microbiology*, 27, 728–736.
- Johnson (2008) Evolutionary Genetics of Hybrid Incompatibility | World Library of Science.
- Jones MB, Highlander SK, Anderson EL *et al.* (2015) Library preparation methodology can influence genomic and functional predictions in human microbiome research. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 14024–14029.
- Kalbe M, Kurtz J (2006) Local differences in immunocompetence reflect resistance of sticklebacks against the eye fluke *Diplostomum pseudospathaceum*. *Parasitology*, 132, 105–116.

- Kalbe M, Wegner KM, Reusch TBH (2002) Dispersion patterns of parasites in 0+ year three-spined sticklebacks: a cross population comparison. *Journal of Fish Biology*, 60, 1529–1542.
- Kaltz O, Shykoff JA (1998) Local adaptation in host-parasite systems. *Heredity*, 81, 361–370.
- Karvonen A, Kristjánsson BK, Skúlason S *et al.* (2013a) Water temperature, not fish morph, determines parasite infections of sympatric Icelandic threespine sticklebacks (*Gasterosteus aculeatus*). *Ecology and Evolution*, 3, 1507–1517.
- Karvonen A, Lucek K, Marques DA, Seehausen O (2015) Divergent Macroparasite Infections in Parapatric Swiss Lake-Stream Pairs of Threespine Stickleback (*Gasterosteus aculeatus*). *PLoS ONE*, 10, e0130579.
- Karvonen A, Lundsgaard-Hansen B, Jokela J, Seehausen O (2013b) Differentiation in parasitism among ecotypes of whitefish segregating along depth gradients. *Oikos*, 122, 122–128.
- Karvonen A, Rintamäki P, Jokela J, Valtonen ET (2010) Increasing water temperature and disease risks in aquatic systems: Climate change increases the risk of some, but not all, diseases. *International Journal for Parasitology*, 40, 1483–1488.
- Karvonen A, Seehausen O (2012) The Role of Parasitism in Adaptive Radiations; When Might Parasites Promote and When Might They Constrain Ecological Speciation? *International Journal of Ecology*, 2012, e280169.
- Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, 7, 1225–1241.
- Kennedy NA, Walker AW, Berry SH *et al.* (2014) The Impact of Different DNA Extraction Kits and Laboratories upon the Assessment of Human Gut Microbiota Composition by 16S rRNA Gene Sequencing. *PLOS ONE*, 9, e88982.

References

- Kim M, Yu Z (2014) Variations in 16S rRNA-based microbiome profiling between pyrosequencing runs and between pyrosequencing facilities. *Journal of Microbiology*, 52, 355–365.
- Kimes NE, Grim CJ, Johnson WR *et al.* (2012) Temperature regulation of virulence factors in the pathogen *Vibrio coralliilyticus*. *The ISME Journal*, 6, 835–846.
- Koch H, Cisarovsky G, Schmid-Hempel P (2012) Ecological effects on gut bacterial communities in wild bumblebee colonies. *Journal of Animal Ecology*, 81, 1202–1210.
- Koli L (1969) Geographical variation of *Cottus gobio* L. (Pisces, Cottidae) in Northern Europe. *Annales Zoologici Fennici*, 6, 353–390.
- Kottelat M, Freyhof J (2007) *Handbook of European freshwater fishes*. Publ. Kottelat, Cornol.
- Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*.
- Kueneman JG, Parfrey LW, Woodhams DC *et al.* (2014) The amphibian skin-associated microbiome across species, space and life history stages. *Molecular Ecology*, 23, 1238–1250.
- Laguerre C, Kaldonski N, Perrot-Minnot MJ, Motreuil S, Bollache L (2007) Modification of Hosts' Behavior by a Parasite: Field Evidence for Adaptive Manipulation. *Ecology*, 88, 2839–2847.
- Lande R (1981) Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences*, 78, 3721–3725.
- Larsen AM, Bullard SA, Womble M, Arias CR (2015) Community Structure of Skin Microbiome of Gulf Killifish, *Fundulus grandis*, Is Driven by Seasonality and Not Exposure to Oiled Sediments in a Louisiana Salt Marsh. *Microbial Ecology*, 70, 534–544.

- Larsen A m., Mohammed H h., Arias C r. (2014a) Characterization of the gut microbiota of three commercially valuable warmwater fish species. *Journal of Applied Microbiology*, 116, 1396–1404.
- Larsen AM, Mohammed HH, Arias CR (2014b) Comparison of DNA extraction protocols for the analysis of gut microbiota in fishes. *FEMS Microbiology Letters*, fnu031.
- Larsen A, Tao Z, Bullard SA, Arias CR (2013) Diversity of the skin microbiota of fishes: evidence for host species specificity. *FEMS Microbiology Ecology*, 85, 483–494.
- Laukens D, Brinkman BM, Raes J, De Vos M, Vandenabeele P (2016) Heterogeneity of the gut microbiome in mice: guidelines for optimizing experimental design. *FEMS Microbiology Reviews*, 40, 117–132.
- Lazado CC, Caipang CMA (2014) Mucosal immunity and probiotics in fish. *Fish & Shellfish Immunology*, 39, 78–89.
- Le Brun N, Renaud F, Berrebi P, Lambert A (1992) Hybrid Zones and Host-Parasite Relationships: Effect on the Evolution of Parasitic Specificity. *Evolution*, 46, 56–61.
- Lexer C, Randell RA, Rieseberg LH (2003) Experimental Hybridization as a Tool for Studying Selection in the Wild. *Ecology*, 84, 1688–1699.
- Ley RE, Bäckhed F, Turnbaugh P *et al.* (2005) Obesity alters gut microbial ecology. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 11070–11075.
- Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI (2008) Worlds within worlds: evolution of the vertebrate gut microbiota. *Nature Reviews Microbiology*, 6, 776–788.
- Linnenbrink M, Wang J, Hardouin EA *et al.* (2013) The role of biogeography in shaping diversity of the intestinal microbiota in house mice. *Molecular Ecology*, 22, 1904–1916.
- Lively CM, Dybdahl MF (2000) Parasite adaptation to locally common host genotypes. *Nature*, 405, 679–681.

References

- Lokmer A, Wegner KM (2015) Hemolymph microbiome of Pacific oysters in response to temperature, temperature stress and infection. *The ISME journal*, 9, 670–682.
- Long AD, Mullaney SL, Reid LA *et al.* (1995) High Resolution Mapping of Genetic Factors Affecting Abdominal Bristle Number in *Drosophila Melanogaster*. *Genetics*, 139, 1273–1291.
- Lowrey L, Woodhams DC, Tacchi L, Salinas I (2015) Topographical mapping of the rainbow trout (*Oncorhynchus mykiss*) microbiome reveals a diverse bacterial community in the skin with antifungal properties. *Applied and Environmental Microbiology*, AEM.01826-15.
- Lowry DB, Modliszewski JL, Wright KM, Wu CA, Willis JH (2008a) The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 363, 3009–3021.
- Lowry DB, Rockwood RC, Willis JH (2008b) Ecological Reproductive Isolation of Coast and Inland Races of *Mimulus Guttatus*. *Evolution*, 62, 2196–2214.
- Lozupone C, Knight R (2005) UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Applied and Environmental Microbiology*, 71, 8228–8235.
- Lozupone CA, Knight R (2007) Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences*, 104, 11436–11440.
- Lozupone CA, Knight R (2008) Species Divergence and the Measurement of Microbial Diversity. *FEMS microbiology reviews*, 32, 557–578.
- Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R (2011) UniFrac: an effective distance metric for microbial community comparison. *The ISME journal*, 5, 169–172.
- Lozupone CA, Stombaugh J, Gonzalez A *et al.* (2013) Meta-analyses of studies of the human microbiota. *Genome Research*, 23, 1704–1714.
- Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R (2012) Diversity, stability and resilience of the human gut microbiota. *Nature*, 489, 220–230.

- Lu G, Bernatchez L (1999) Correlated Trophic Specialization and Genetic Divergence in Sympatric Lake Whitefish Ecotypes (*Coregonus clupeaformis*): Support for the Ecological Speciation Hypothesis. *Evolution*, 53, 1491–1505.
- Maan ME, Seehausen O (2011) Ecology, sexual selection and speciation. *Ecology Letters*, 14, 591–602.
- Maan ME, Van Rooijen AMC, Van Alphen JJM, Seehausen O (2008) Parasite-mediated sexual selection and species divergence in Lake Victoria cichlid fish. *Biological Journal of the Linnean Society*, 94, 53–60.
- MacColl ADC, Chapman SM (2010) Parasites can cause selection against migrants following dispersal between environments. *Functional Ecology*, 24, 847–856.
- Macnab V, Barber I (2012) Some (worms) like it hot: fish parasites grow faster in warmer water, and alter host thermal preferences. *Global Change Biology*, 18, 1540–1548.
- Macpherson AJ, Harris NL (2004) Interactions between commensal intestinal bacteria and the immune system. *Nature Reviews Immunology*, 4, 478–485.
- Magurran AE (1988) *Ecological Diversity and Its Measurement*. Springer, Dordrecht.
- Maheshwari S, Barbash DA (2011) The Genetics of Hybrid Incompatibilities. *Annual Review of Genetics*, 45, 331–355.
- Mallet J (1995) A species definition for the modern synthesis. *Trends in Ecology & Evolution*, 10, 294–299.
- Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, 20, 229–237.
- Mallet J (2007) Hybrid speciation. *Nature*, 446, 279–283.
- Mallet J (2008) Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 363, 2971–2986.

References

- Mallet J, Barton NH (1989) Strong Natural Selection in a Warning-Color Hybrid Zone. *Evolution*, 43, 421–431.
- Mani GS, Clarke BC (1990) Mutational Order: A Major Stochastic Process in Evolution. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 240, 29–37.
- Mao D-P, Zhou Q, Chen C-Y, Quan Z-X (2012) Coverage evaluation of universal bacterial primers using the metagenomic datasets. *BMC Microbiology*, 12, 1–8.
- Margulis L (1993) *Symbiosis in cell evolution: microbial communities in the Archean and Proterozoic eons*. Freeman, New York.
- Mayr E (1942) *Systematics and the Origin of Species*. Columbia University Press, New York.
- McFall-Ngai M, Hadfield MG, Bosch TCG *et al.* (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, 3229–3236.
- Mizrahi-Man O, Davenport ER, Gilad Y (2013) Taxonomic Classification of Bacterial 16S rRNA Genes Using Short Sequencing Reads: Evaluation of Effective Study Designs. *PLoS ONE*, 8, e53608.
- Montgomery MK, McFall-Ngai M (1994) Bacterial symbionts induce host organ morphogenesis during early postembryonic development of the squid *Euprymna scolopes*. *Development*, 120, 1719–1729.
- Moran NA (2006) Symbiosis. *Current Biology*, 16, R866–R871.
- Moravec F (2001) Common sculpin *Cottus gobio* as a natural paratenic host of *Proteocephalus longicollis* (Cestoda: Proteocephalidae), a parasite of salmonids, in Europe. *Diseases of aquatic organisms*, 45, 155–158.
- Moullia C (1999) Parasitism of Plant and Animal Hybrids: Are Facts and Fates the same? *Ecology*, 80, 392–406.
- Muzzall PM, Bowen CA (2002) Parasites of the Slimy Sculpin, *Cottus cognatus* Richardson, 1836, from Lake Huron, U.S.A. *Comparative Parasitology*, 69, 196–201.

- Neuman C, Hatje E, Zarkasi KZ *et al.* (2016) The effect of diet and environmental temperature on the faecal microbiota of farmed Tasmanian Atlantic Salmon (*Salmo salar* L.). *Aquaculture Research*, 47, 660–672.
- Niewiadomska K (1984) Present status of *Diplostomum spathaceum* (Rudolphi, 1819) and differentiation of *Diplostomum pseudospathaceum* nom. nov. (Trematoda: Diplostomatidae). *Systematic Parasitology*, 6, 81–86.
- Nolte AW, Freyhof J, Stemshorn KC, Tautz D (2005) An invasive lineage of sculpins, *Cottus* sp. (Pisces, Teleostei) in the Rhine with new habitat adaptations has originated from hybridization between old phylogeographic groups. *Proceedings of the Royal Society B: Biological Sciences*, 272, 2379–2387.
- Nolte AW, Freyhof J, Tautz D (2006) When invaders meet locally adapted types: rapid moulding of hybrid zones between sculpins (*Cottus*, Pisces) in the Rhine system. *Molecular Ecology*, 15, 1983–1993.
- Nolte a. W, Gompert Z, Buerkle CA (2009) Variable patterns of introgression in two sculpin hybrid zones suggest that genomic isolation differs among populations. *Molecular Ecology*, 18, 2615–2627.
- Nolte AW, Sheets HD (2005) Shape based assignment tests suggest transgressive phenotypes in natural sculpin hybrids (Teleostei, Scorpaeniformes, Cottidae). *Frontiers in Zoology*, 2, 11.
- Nolte AW, Tautz D (2010) Understanding the onset of hybrid speciation. *Trends in Genetics: TIG*, 26, 54–58.
- Noor MAF, Feder JL (2006) Speciation genetics: evolving approaches. *Nature Reviews Genetics*, 7, 851–861.
- Nosil P (2012) *Ecological speciation*. Oxford University Press, Oxford ; New York.
- Nosil P, Flaxman SM (2011) Conditions for mutation-order speciation. *Proceedings of the Royal Society B: Biological Sciences*, 278, 399–407.

References

- Nosil P, Schluter D (2011) The genes underlying the process of speciation. *Trends in Ecology & Evolution*, 26, 160–167.
- Nosil P, Vines TH, Funk DJ (2005) Reproductive Isolation Caused by Natural Selection Against Immigrants from Divergent Habitats. *Evolution*, 59, 705–719.
- Nyholm SV, McFall-Ngai M (2004) The winnowing: establishing the squid–vibrio symbiosis. *Nature Reviews Microbiology*, 2, 632–642.
- Oksanen J, Blanchet, F. Guillaume, Kindt, Roeland *et al.* (2015) vegan: Community Ecology Package.
- Pallares LF, Carbonetto P, Gopalakrishnan S *et al.* (2015) Mapping of Craniofacial Traits in Outbred Mice Identifies Major Developmental Genes Involved in Shape Determination. *PLOS Genet*, 11, e1005607.
- Penders J, Thijs C, Vink C *et al.* (2006) Factors Influencing the Composition of the Intestinal Microbiota in Early Infancy. *Pediatrics*, 118, 511–521.
- Phelan PE, Mellon MT, Kim CH (2005) Functional characterization of full-length TLR3, IRAK-4, and TRAF6 in zebrafish (*Danio rerio*). *Molecular Immunology*, 42, 1057–1071.
- Poulin R (1996) How many parasite species are there: Are we close to answers? *International Journal for Parasitology*, 26, 1127–1129.
- Poulin R (2007) *Evolutionary ecology of parasites*. Princeton University Press.
- Price MN, Dehal PS, Arkin AP (2009) FastTree: Computing Large Minimum Evolution Trees with Profiles instead of a Distance Matrix. *Molecular Biology and Evolution*, 26, 1641–1650.
- Price MN, Dehal PS, Arkin AP (2010) FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE*, 5.
- Price P, Westoby M, Rice B *et al.* (1986) Parasite Mediation in Ecological Interactions. *Annual Review of Ecology and Systematics*, 17, 487–505.

- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multi-locus genotype data. *Genetics*, 155, 945–959.
- Pruesse E, Quast C, Knittel K *et al.* (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research*, 35, 7188–7196.
- Quast C, Pruesse E, Yilmaz P *et al.* (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 41, D590–D596.
- Rabosky DL, Santini F, Eastman J *et al.* (2013) Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications*, 4, 1958.
- R Core Team (2015) R: A Language and Environment for Statistical Computing.
- Rawls JF, Mahowald MA, Ley RE, Gordon JI (2006) Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. *Cell*, 127, 423–433.
- Rawls JF, Samuel BS, Gordon JI (2004) Gnotobiotic zebrafish reveal evolutionarily conserved responses to the gut microbiota. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 4596–4601.
- Reusch TBH, Häberli MA, Aeschlimann PB, Milinski M (2001) Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature*, 414, 300–302.
- Rieseberg LH, Archer MA, Wayne RK (1999a) Transgressive segregation, adaptation and speciation. *Heredity*, 83, 363–372.
- Rieseberg LH, Buerkle CA (2002) Genetic Mapping in Hybrid Zones. *The American Naturalist*, 159, S36–S50.

- Rieseberg LH, Raymond O, Rosenthal DM *et al.* (2003a) Major Ecological Transitions in Wild Sunflowers Facilitated by Hybridization. *Science*, 301, 1211–1216.
- Rieseberg LH, Sinervo B, Linder CR, Ungerer MC, Arias DM (1996) Role of Gene Interactions in Hybrid Speciation: Evidence from Ancient and Experimental Hybrids. *Science*, 272, 741–745.
- Rieseberg LH, Whitton J, Gardner K (1999b) Hybrid Zones and the Genetic Architecture of a Barrier to Gene Flow Between Two Sunflower Species. *Genetics*, 152, 713–727.
- Rieseberg LH, Widmer A, Arntz AM, Burke B (2003b) The genetic architecture necessary for transgressive segregation is common in both natural and domesticated populations. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 358, 1141–1147.
- Roeselers G, Mittge EK, Stephens WZ *et al.* (2011) Evidence for a core gut microbiota in the zebrafish. *The ISME Journal*, 5, 1595–1608.
- Round JL, Mazmanian SK (2009) The gut microbiota shapes intestinal immune responses during health and disease. *Nature Reviews Immunology*, 9, 313–323.
- Rubin BER, Sanders JG, Hampton-Marcell J *et al.* (2014) DNA extraction protocols cause differences in 16S rRNA amplicon sequencing efficiency but not in community profile composition or structure. *Microbiologyopen*, 3, 910–921.
- Rumpus AE (1975) The helminth parasites of the bullhead *Cottus gobio* (L.) and the stone loach *Noemacheilus barbatulus* (L.) from the River Avon, Hampshire. *Journal of Fish Biology*, 7, 469–483.
- Rundle HD (2002) A test of ecologically dependent postmating isolation between sympatric sticklebacks. *Evolution*, 56, 322–329.
- Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters*, 8, 336–352.
- Sadier A, Viriot L, Pantalacci S, Laudet V (2014) The ectodysplasin pathway: from diseases to adaptations. *Trends in Genetics*, 30, 24–31.

- Sage RD, Heyneman D, Lim K-C, Wilson AC (1986) Wormy mice in a hybrid zone. *Nature*, 324, 60–63.
- Salinas I, Zhang Y-A, Sunyer JO (2011) Mucosal immunoglobulins and B cells of teleost fish. *Developmental & Comparative Immunology*, 35, 1346–1365.
- Santini F, Miglietta MP, Faucci A (2012) Speciation: Where Are We Now? An Introduction to a Special Issue on Speciation. *Evolutionary Biology*, 39, 141–147.
- Sar N, Rosenberg E (1987) Fish skin bacteria: Colonial and cellular hydrophobicity. *Microbial Ecology*, 13, 193–202.
- Schade FM, Raupach MJ, Mathias Wegner K (2015) Seasonal variation in parasite infection patterns of marine fish species from the Northern Wadden Sea in relation to inter-annual temperature fluctuations. *Journal of Sea Research*.
- Scharsack JP, Kalbe M (2014) Differences in susceptibility and immune responses of three-spined sticklebacks (*Gasterosteus aculeatus*) from lake and river ecotypes to sequential infections with the eye fluke *Diplostomum pseudospathaceum*. *Parasites & Vectors*, 7, 109.
- Scharsack JP, Kalbe M, Harrod C, Rauch G (2007) Habitat-specific adaptation of immune responses of stickleback (*Gasterosteus aculeatus*) lake and river ecotypes. *Proceedings of the Royal Society B: Biological Sciences*, 274, 1523–1532.
- Schloss PD (2009) A High-Throughput DNA Sequence Aligner for Microbial Ecology Studies. *PLoS ONE*, 4, e8230.
- Schloss PD, Gevers D, Westcott SL (2011) Reducing the Effects of PCR Amplification and Sequencing Artifacts on 16S rRNA-Based Studies. *PLoS ONE*, 6, e27310.
- Schluter D (2000) Ecological character displacement in adaptive radiation. *American Naturalist*, 156, S4–S16.
- Schluter D (2001) Ecology and the origin of species. *Trends in Ecology & Evolution*, 16, 372–380.

- Schluter D (2009) Evidence for Ecological Speciation and Its Alternative. *Science*, 323, 737–741.
- Schmitt-Wagner D, Friedrich MW, Wagner B, Brune A (2003) Phylogenetic Diversity, Abundance, and Axial Distribution of Bacteria in the Intestinal Tract of Two Soil-Feeding Termites (*Cubitermes* spp.). *Applied and Environmental Microbiology*, 69, 6007–6017.
- Schwartz AK, Hendry AP (2007) A test for the parallel co-evolution of male colour and female preference in Trinidadian guppies (*Poecilia reticulata*). *Evolutionary Ecology Research*, 9.
- Seehausen O (2004) Hybridization and adaptive radiation. *Trends in Ecology & Evolution*, 19, 198–207.
- Seehausen O, Butlin RK, Keller I *et al.* (2014) Genomics and the origin of species. *Nature Reviews Genetics*, 15, 176–192.
- Seehausen O, Wagner CE (2014) Speciation in Freshwater Fishes. *Annual Review of Ecology, Evolution, and Systematics*, 45, 621–651.
- Sekirov I, Russell SL, Antunes LCM, Finlay BB (2010) Gut Microbiota in Health and Disease. *Physiological Reviews*, 90, 859–904.
- Seppälä O, Karvonen A, Valtonen ET (2011) Eye fluke-induced cataracts in natural fish populations: is there potential for host manipulation? *Parasitology*, 138, 209–214.
- Šimková A, Dávidová M, Papoušek I, Vetešník L (2013) Does interspecies hybridization affect the host specificity of parasites in cyprinid fish? *Parasites & Vectors*, 6, 95.
- Song SJ, Lauber C, Costello EK *et al.* (2013) Cohabiting family members share microbiota with one another and with their dogs. *eLife*, 2.
- Spor A, Koren O, Ley R (2011) Unravelling the effects of the environment and host genotype on the gut microbiome. *Nature Reviews Microbiology*, 9, 279–290.

- Srinivas G, Moller S, Wang J *et al.* (2013) Genome-wide mapping of gene-microbiota interactions in susceptibility to autoimmune skin blistering. *Nature Communications*, 4.
- Staubach F, Baines JF, Künzel S, Bik EM, Petrov DA (2013) Host Species and Environmental Effects on Bacterial Communities Associated with *Drosophila* in the Laboratory and in the Natural Environment. *PLoS ONE*, 8, e70749.
- Stemshorn K (2007) The genomic Make-Up of a Hybrid Species - Analysis of the Invasive Cottus Lineage (Pisces, Teleostei) in the River Rhine system. Doctoral Thesis Thesis. Universität zu Köln.
- Stemshorn KC, Reed FA, Nolte AW, Tautz D (2011) Rapid formation of distinct hybrid lineages after secondary contact of two fish species (*Cottus* sp.). *Molecular Ecology*, 20, 1475–1491.
- Sullam KE, Essinger SD, Lozupone CA *et al.* (2012) Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. *Molecular Ecology*, 21, 3363–3378.
- The Human Microbiome Project Consortium (2012) Structure, function and diversity of the healthy human microbiome. *Nature*, 486, 207–214.
- Tlaskalová-Hogenová H, Štěpánková R, Hudcovic T *et al.* (2004) Commensal bacteria (normal microflora), mucosal immunity and chronic inflammatory and autoimmune diseases. *Immunology Letters*, 93, 97–108.
- Tsuchida T, Koga R, Fukatsu T (2004) Host Plant Specialization Governed by Facultative Symbiont. *Science*, 303, 1989–1989.
- Turnbaugh PJ, Hamady M, Yatsunenko T *et al.* (2009) A core gut microbiome in obese and lean twins. *Nature*, 457, 480–484.
- Turner LM, Harr B (2014) Genome-wide mapping in a house mouse hybrid zone reveals hybrid sterility loci and Dobzhansky-Muller interactions. *Elife*, 3, e02504.

References

- Vale PF, Salvaudon L, Kaltz O, Fellous S (2008) The role of the environment in the evolutionary ecology of host parasite interactions - Meeting report, Paris, 5th December, 2007. *Infection Genetics and Evolution*, 8, 302–305.
- van Valen L (1973) A New Evolutionary Law. *Evolutionary Theory*, 1, 1-30.
- Via S, Bouck AC, Skillman S (2000) Reproductive isolation between divergent races of pea aphids on two hosts. ii. selection against migrants and hybrids in the parental environments. *Evolution*, 54, 1626–1637.
- Vijay-Kumar M, Aitken JD, Carvalho FA *et al.* (2010) Metabolic Syndrome and Altered Gut Microbiota in Mice Lacking Toll-Like Receptor 5. *Science*, 328, 228–231.
- Vojtek J (1964) Zur Kenntnis des Entwicklungszyklus von *Apatemon cobitidis* (Linstow, 1890). *Zeitschrift für Parasitenkunde*, 24, 578–599.
- Volckaert F a. M, Hänfling B, Hellemans B, Carvalho GR (2002) Timing of the population dynamics of bullhead *Cottus gobio* (Teleostei: Cottidae) during the Pleistocene. *Journal of Evolutionary Biology*, 15, 930–944.
- Walke JB, Becker MH, Loftus SC *et al.* (2014) Amphibian skin may select for rare environmental microbes. *The ISME Journal*, 8, 2207–2217.
- Walker AW, Duncan SH, Leitch ECM, Child MW, Flint HJ (2005) pH and Peptide Supply Can Radically Alter Bacterial Populations and Short-Chain Fatty Acid Ratios within Microbial Communities from the Human Colon. *Applied and Environmental Microbiology*, 71, 3692–3700.
- Wallace AR (1869) *The Malay Archipelago: The Land of the Orang-Utan and the Bird of Paradise. A Narrative of Travel, with Studies of Man and Nature*. Macmillan.
- Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology*, 73, 5261–5267.

- Wang J, Kalyan S, Steck N *et al.* (2015) Analysis of intestinal microbiota in hybrid house mice reveals evolutionary divergence in a vertebrate hologenome. *Nature Communications*, 6.
- Wang W, Zhou Z, He S *et al.* (2010) Identification of the adherent microbiota on the gills and skin of poly-cultured gibel carp (*Carassius auratus gibelio*) and bluntnose black bream (*Megalobrama amblycephala* Yih). *Aquaculture Research*, 41, e72–e83.
- Warnecke F, Luginbühl P, Ivanova N *et al.* (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature*, 450, 560–565.
- Wegner KM, Reusch TBH, Kalbe M (2003) Multiple parasites are driving major histocompatibility complex polymorphism in the wild. *Journal of Evolutionary Biology*, 16, 224–232.
- Wiley C, Qvarnström A, Gustafsson L (2009) Effects of hybridization on the immunity of collared Ficedula albicollis and pied flycatchers F. hypoleuca, and their infection by haemosporidians. *Journal of Avian Biology*, 40, 352–357.
- Windsor DA (1998) Controversies in parasitology, Most of the species on Earth are parasites. *International Journal for Parasitology*, 28, 1939–1941.
- Winger AC, Hansen H, Bachmann L, Bakke TA (2008) Gyrodactylus species (Monogenea) infecting alpine bullhead (Cottus poecilopus Heckel, 1837) in Norway and Slovakia, including the description of Gyrodactylus mariannae sp. nov. *Acta Parasitologica*, 53, 240–250.
- Wolf JBW, Lindell J, Backström N (2010) Speciation genetics: current status and evolving approaches. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365, 1717–1733.
- Wolinska J, Lively CM, Spaak P (2008) Parasites in hybridizing communities: the Red Queen again? *Trends in Parasitology*, 24, 121–126.

References

- Wong S, Rawls JF (2012) Intestinal microbiota composition in fishes is influenced by host ecology and environment. *Molecular Ecology*, 21, 3100–3102.
- Wood CL, Byers JE, Cottingham KL *et al.* (2007) Parasites alter community structure. *Proceedings of the National Academy of Sciences*, 104, 9335–9339.
- Wood TE, Takebayashi N, Barker MS *et al.* (2009) The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences*, 106, 13875–13879.
- Wu C-I, Ting C-T (2004) Genes and speciation. *Nature Reviews Genetics*, 5, 114–122.
- Yatsunencko T, Rey FE, Manary MJ *et al.* (2012) Human gut microbiome viewed across age and geography. *Nature*, 486, 222–227.
- Ye L, Amberg J, Chapman D, Gaikowski M, Liu W-T (2014) Fish gut microbiota analysis differentiates physiology and behavior of invasive Asian carp and indigenous American fish. *The ISME Journal*, 8, 541–551.
- Yuan S, Cohen DB, Ravel J, Abdo Z, Forney LJ (2012) Evaluation of Methods for the Extraction and Purification of DNA from the Human Microbiome. *PLOS ONE*, 7, e33865.
- Yun J-H, Roh SW, Whon TW *et al.* (2014) Insect Gut Bacterial Diversity Determined by Environmental Habitat, Diet, Developmental Stage, and Phylogeny of Host. *Applied and Environmental Microbiology*, 80, 5254–5264.
- Zoetendal EG, Akkermans ADL, Akkermans-van Vliet WM (2001) The Host Genotype Affects the Bacterial Community in the Human Gastrointestinal Tract. *Microbial Ecology in Health and Disease*, 13, 129–134.
- Zuur A, Ieno EM, Smith GM (2007) *Analyzing Ecological Data*. Springer Science + Business Media, LLC, New York.

DANKSAGUNG

Ohne die Hilfe und die moralische und tatkräftige Unterstützung vieler lieber Freunde, Kollegen und Familienmitglieder wäre diese Arbeit so nicht möglich gewesen, deshalb nutze ich hiermit die Gelegenheit allen zu danken.

Zuallererst möchte ich meinem Betreuer Arne Nolte danken, dass er mich in seine Gruppe aufgenommen und mir ermöglicht hat zu promovieren und mich zur Wissenschaftlerin zu entwickeln. Danke für jegliche Unterstützung und die ansteckende Begeisterung für Evolutionsbiologie im Allgemeinen, Fische im Besonderen, und für alles was da sonst so „krecht und fleucht“. Du hast stets ein offenes Ohr gehabt, meine Bedenken ernst genommen und mir Mut gemacht. Danke auch an Diethard Tautz, dass er mich in seinem Department aufgenommen und stets gerne über verschiedenste Themen zu Evolutionsbiologie diskutiert hat.

Ich danke allen ehemaligen und jetzigen Mitgliedern der *Cottus*-Gruppe. Danke an Till Cypionka und Jie Cheng, die mit mir ihr Büro geteilt haben und mir nicht nur fachlich zur Seite standen. Danke an Till für die vielen Späße und Danke an Jie, dass ich so viel über China von dir lernen durfte. Stefan Dennenmoser danke ich für den Austausch über Fotografie. Danke an Ela Iwaskiewitz für die tolle Zeit im Büro und die Schokolade. Danke auch an Elke Bustorf, die nicht nur immer ein offenes Ohr für alle laborrelevanten Fragen hatte, sondern mich auch bei allen anderen Problemen unterstützt hat. Danke auch für die tollen Blumen zu meinen Geburtstagen. Vielen Dank an Heike Harre, die sich die ersten Jahre halbtags um unsere Groppen gekümmert hat. Danke auch für die netten Pausen und dein stetes Interesse an meiner Arbeit. Und nicht zuletzt Danke an Ralf, ohne dich wären die Fische nicht halb so moppelig und zufrieden wie sie es jetzt sind. Danke auch für die Fahrten zu Hof Biss und zum Theresienhof.

Ohne die bereitwillige Unterstützung von Wilhelm Kreuzmann, dem Geschäftsführer der Sieg Fischerei-Genossenschaft, und der tatkräftigen Hilfe von Fabian Herder (Zoologisches Forschungsmuseum Alexander Koenig, Bonn) hätten wir nicht so viele Groppen, Bachschmerlen und Elritzen für mein Projekt fangen können. Danke für die spannenden Tage an der Sieg und das leckere Abendessen.

Besonderer Dank gilt John Baines und seiner Gruppe. Er hat nicht nur wesentlich zur Projektidee beigetragen sondern mir auch einen tollen Laborplatz zur Verfügung gestellt. Außerdem hat er mir als Mitglied in meinem thesis committee all die Jahre beratend zur Seite gestanden. Natürlich möchte ich allen Mitgliedern der Baines-Gruppe danken. Silke Carstensen, danke für das Anleiten im Labor und die stete Hilfe, auch dafür danke an Jan Schubert. Marie Vallier und Meriem Belheouane waren nicht nur in laborrelevanten Fragen wichtige Ansprechpartnerinnen für mich sondern haben mir auch unzählige Fragen zur Sequenzanalyse beantwortet. Danke auch an Philipp Rausch und Jun Wang für die Hilfe bei Analyse- und Statistikfragen. Sven Künzel hat nicht nur beim Sequenzieren meiner Proben tolle Arbeit geleistet sondern hat mich immer aufgebaut und ermutigt, wenn etwas im Labor nicht so recht klappen wollte (ich sage nur „Labor-Voodoo“). Vielen Dank dafür.

Danksagung

Vielen Dank auch an Martin Kalbe für die Unterstützung und den Rat als thesis committee Mitglied. Danke für die Einführung in die spannende Welt der Fischparasiten und die große Hilfe mit den Parasitenprojekten. Lieben Dank auch an Nicole Thomsen, die mich tatkräftig und gewissenhaft bei der Vorbereitung der Sequenzierproben unterstützt hat. An Kerstin Mehnert vielen Dank für ihr Engagement als Koordinatorin der IMPRS. Du standest mir stets mit Rat und Tat zur Seite und hast geduldig unzählige Fragen beantwortet. Danke an Alina Jeschke, du hast mir mit Ela die letzten Monate eine tolle Zeit im Büro beschert, ich werde euch beide sehr vermissen. Danke an alle PhD-Studenten meiner Kohorte: Luisa, Meriem, Anne W., Anne B., Ozan and Vienna. Wir hatten eine tolle Zeit und es war immer schön euch zu sehen. Natürlich danke auch an unzählige Kollegen, mit denen ich ernste und lustige Gespräche (und viele Gespräche über Essen) während der Zugfahrten hatte. Lieben Dank auch an Heike Gruber, bei der ich das letzte halbe Jahr immer mit nach Plön fahren durfte.

Danke an das tollste IT-Team: Derk Wachsmuth, Werner Wegner und Herbert Kieseewetter. Ihr habt so manches PC-Problem umgehend gelöst und ward auch kurz vor Dienstschluss immer geduldig und hilfsbereit. Danke an Iben Martinsen für die schnelle Hilfe beim Bestellen von papern und beim Drucken meiner Arbeit. Außerdem danke an die eher unsichtbaren Helferlein vom Hausmeister- und Raumpflege-Team. Ohne euch wären wir alle aufgeschmissen. Vielen Dank auch an die Verwaltung und Frau Hammerich für die Hilfe bei Reiseanträgen, Buchungen und vielem mehr.

Großer Dank geht an Miriam Linnenbrink, die im Laufe der Jahre zu einer guten Freundin geworden ist. Danke für's Ermutigen, für's Zuhören, für all die lustigen Stunden und die unzähligen gemeinsamen Mittagspausen. Vielen Dank auch an Natascha Hasenkamp, die mich erst als Gleichgesinnte und dann als Koordinatorin immer unterstützt hat. Danke für die vielen vertratschten Pausen und die gemeinsamen Zugfahrten.

Ganz besonders danken möchte ich Vienna Kowallik und Philipp Dirksen, die mich seit dem 1. Semester Biologie begleiten und zu den mir wichtigsten Menschen geworden sind. Danke für die fachliche und moralische Unterstützung, danke für's Zuhören, für's Aufbauen, für's Mutmachen, für's Ablenken, für's Korrekturlesen und noch viel viel mehr, was ich gar nicht alles aufschreiben kann. Danke auch an alle anderen Freunde für euer Interesse, euer Verständnis, wenn ich mal wieder keine Zeit hatte, und für die tolle Ablenkung durch Koch- Spiele- und Plauderabende.

Meinen Eltern und auch meinen Brüdern, meiner Oma und dem Rest der Familie möchte ich dafür danken, dass ihr immer an mich und meine Arbeit geglaubt habt. Ohne euch wäre diese Dissertation nicht möglich gewesen. Oma, wenn du das liest, werde ich endlich bald meinen Doktorhut bekommen. Sinha, danke für deine unendliche Geduld mit der du alle Höhen und Tiefen meiner Arbeit mitgemacht hast. Danke auch für das Bekochen und das stete Mutmachen und dafür, dass du immer für mich da bist.

AFFIDAVIT

Hereby I declare that this dissertation, in form and content and except for advices given by my supervisor, Prof. Dr. Arne W. Nolte, constitutes my own work. None of the chapters have been published before. This work neither in part nor as a whole has not been subject to any examination procedure somewhere else before. This dissertation was prepared in compliance with "Rules of Good Scientific Practice" of the German Research Foundations.

Kiel, den

Sunna Ellendt