Patterns of Hybridization Among Catostomids of the Upper Colorado River Basin and the Diversity of Their Associated Microbiota

Se Jin Song
University of Colorado at Boulder, sejin.song@colorado.edu

Follow this and additional works at: https://scholar.colorado.edu/ebio_gradetds

Part of the Environmental Studies Commons, Evolution Commons, and the Microbiology Commons

Recommended Citation
https://scholar.colorado.edu/ebio_gradetds/48

This Dissertation is brought to you for free and open access by Ecology & Evolutionary Biology at CU Scholar. It has been accepted for inclusion in Ecology & Evolutionary Biology Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
Patterns of Hybridization Among Catostomids of the Upper Colorado River Basin

and the Diversity of Their Associated Microbiota

by

Se Jin Song

B.S., University of Georgia, 2004

A thesis submitted to the

Faculty of the Graduate School of the University of Colorado

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Ecology and Evolutionary Biology

2013
This thesis entitled: Patterns of Hybridization Among Catostomids of the Upper Colorado River Basin and the Diversity of Their Associated Microbiota written by Se Jin Song has been approved for the Department of Ecology and Evolutionary Biology.

Jeffry B. Mitton

Rob Knight

Date___________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.
Song, Se Jin (Ph.D., Ecology and Evolutionary Biology)

Patterns of hybridization among catostomids of the Upper Colorado River Basin and the diversity of their associated microbiota

Thesis directed by Professor Jeffry B. Mitton

A major tenet of evolutionary biology is that the fitness of an individual is determined by its genetics and its interactions with the environment. Yet microbial ecology has taught us that the microbes with which animals live in close association also constitute an important component of the animal's phenotype, playing necessary roles in their development, nutrition, and immunity; thus, they too can affect animal fitness and their interactions with the environment. Consequently, a better understanding of both of these factors is essential, and has significant implications for the conservation of natural systems. My dissertation research examines these two factors and thus lies at the interface of evolutionary genetics and microbial ecology by investigating the consequences of hybridization among three catostomid fish species and analyzing the factors affecting the diversity of their associated microbes. Two of these species are native to the Upper Colorado River Basin, and an introduced third species has spread throughout the ranges of the two native fish.

Although it is well known that non-native fish introductions often cause the decline of native species, the effects of subsequent hybridization on the potential survival of the native species are often challenging to quantify. To meet this challenge, I surveyed the genetic composition of hybrid individuals among all three species across a wide geographic range and found that the bulk of hybridization occurs between the introduced sucker and
one of the native species and that this process is slowly destroying the genetic integrity of
the native suckers. Moreover, the range of hybrid types found was related to the amount of
phylogenetic distance and overlap of biogeographic history among species.

Second, I investigated the diversity of microbes associated with the gastrointestinal
tracts and epidermal mucus of these catostomids and my findings indicate that while host
diet and possibly host phylogeny play a role in shaping their gut microbial communities, the
environment plays a larger role than these factors in shaping their skin microbiota.
Collectively, this dissertation highlights how evolutionary history can determine the
potential for hybridization between species and interacts with the environment to influence
host-microbe relationships.
ACKNOWLEDGEMENTS

First and foremost, I thank my graduate advisor, Jeff Mitton, who has not only been an incredible academic mentor, but also my biggest supporter. He always encouraged me to pursue my interests, no matter how broad, and his enthusiasm never waned. Thanks to Andy Martin, Rob Guralnick, Patrik Nosil, and Rob Knight for their service on my dissertation committee. Their invaluable insight, influence and support made this work what it is and more importantly, they have been instrumental in my academic, professional, and personal development. I am also grateful to Tom Parchman and Dave McDonald for providing their expertise on sucker genetics.

Much of this work was done in collaboration with Colorado Parks and Wildlife, Utah Division of Wildlife Resources, and the U.S. Fish and Wildlife Service, and this research would not have been possible without the funds provided by the University of Colorado Graduate School, the University of Colorado Museum, the University of Colorado Department of Ecology and Evolutionary Biology, and the Earth Microbiome Project.

Throughout my graduate career, I am fortunate to have had a wonderful group of colleagues who I can also call friends, especially my cohort. These are all people I was able to commiserate, debate, and celebrate with through this long and arduous journey, and I am grateful that so many of them will continue to be my collaborators in the future. Thanks to my supportive and loving family, especially my mom, who despite never having understood my work, has always shown the greatest confidence in me and has pushed me to never stop reaching for my goals, whatever they may be. And finally, a special thanks to my amazing husband, Ian, for his unwavering love and support, for being my constant and for always reminding me of the things that I love most in this world.
Table of Contents

ACKNOWLEDGEMENTS.................................................................................................................. v

TABLES .......................................................................................................................................... vii

FIGURES ......................................................................................................................................... viii

CHAPTER 1: Introduction .............................................................................................................. 1

INTRODUCTION ................................................................................................................................. 8

MATERIALS AND METHODS ........................................................................................................... 12

RESULTS ......................................................................................................................................... 18

DISCUSSION ................................................................................................................................. 28

CHAPTER 2: An analysis of patterns of hybridization between the introduced white sucker
(Catostomus commersoni) and native bluehead (C. discobolus) and flannelmouth (C. latipinnis) suckers in the Upper Colorado River Basin .............................................................. 8

INTRODUCTION ................................................................................................................................. 8

MATERIALS AND METHODS ........................................................................................................... 12

RESULTS ......................................................................................................................................... 18

DISCUSSION ................................................................................................................................. 28

CHAPTER 3: The effects of species traits and environmental factors on the gut and skin
microbiota of wild catostomid fish .................................................................................................. 38

INTRODUCTION ................................................................................................................................. 38

MATERIALS AND METHODS ........................................................................................................... 44

RESULTS ......................................................................................................................................... 50

DISCUSSION ................................................................................................................................. 64

CONCLUSION ................................................................................................................................ 71

CONCLUSION ................................................................................................................................ 73

BIBLIOGRAPHY ............................................................................................................................. 78
### TABLES

**Table 2.1.** Sampling locations and sample sizes by species. Numbers correspond to the locations shown in Figure 2.1. BHS=Bluehead, FMS=Flannelmouth, WHS=White, BxW=Bluehead x White hybrids, FxW=Flannelmouth x White hybrids, BxF=Bluehead x Flannelmouth hybrids. .............................................................. 14

**Table 2.2.** Summary of the genetic diversity found in each population sampled listed by species (Species) and location (Pop). Genetic diversity measures are provided as the percentage of AFLP loci that were polymorphic (%poly loci) as well as expected heterozygosity (Hj), which is also provided with standard error (s.e.(Hj)) and variance (var(Hj)). See Table 2.1 for abbreviations of species and localities. ............................................. 20

**Table 3.1.** Sample types and numbers of each species sampled at each location. Sample sizes of adults (A) and juveniles (J) of each species are given in parentheses. BHS=bluehead sucker, FMS=flannelmouth sucker, WHS=white sucker. ........................................................................................................ 46

**Table 3.2.** Results of PERMANOVA describing how variation in distances is attributed to different factors. Factors of significant effect (<0.05) are bolded. ................. 56

**Table 3.3.** Results of an analysis of similarity (ANOSIM) for significant grouping factors of fecal and skin microbial communities. Only those comparisons with sample sizes high enough for statistical testing are included. Groupings that show significant separation (R>0.25 and p<0.05) are bolded. For those comparisons showing significant grouping (p<0.05), the results of a Levene test of equality of variances are shown. Values from the same analysis using weighted (wtd) Unifrac measures are shown in parentheses. BHS=bluehead sucker, FMS=flannelmouth sucker, WHS=white sucker. COL=Colorado River open site, RFL=Redlands Fish Ladder, GH=Government Highline Fish Ladder. .......... 57

**Table 3.4.** Summary of the best fit linear model for predicting the bacterial diversity of fecal and skin communities. The response of phylogenetic diversity (PD) was modeled using host species identity (Species), age status of the host (AgeGroup), and the location of sampling (Collection Site). ............................................................................. 58

**Table 3.5.** Results from Mantel tests of correlation between distance matrices (DM) generated from host genotype (aflp) and microbial sequence data (fecal or skin). 1000 permutations were performed for each analysis to calculate the significance value (p). The numbers of individuals for which data of both types were available are shown (n). COL=Colorado River open site, RFL=Redlands Fish Ladder site. ........................................................................................................ 64
FIGURES

Figure 2.1. Sampling locations in the Upper Colorado River Basin and the species found at those sites. One location outside of the UCRB (Monument Creek in Colorado Springs, CO, represented in yellow) was sampled. Numbers correspond to the locations listed in Table 2.1. The green star indicates the location of the Redlands Fish Ladder, a fish passage structure located on the Gunnison River near its confluence with the Colorado River. This location was not sampled, but survey data from this locality were used. ................................................................. 13

Figure 2.2. Genetic cluster assignments from an analysis with Structure of 397 individuals into three groups (K=3) based on 457 AFLP loci. Each vertical bar on the graph corresponds to an individual and the different colors represent the proportional membership in each of the three clusters (Blue=BHS, Red=FMS, Yellow=WHS). Individuals are grouped by sampling locations (Colorado=Colorado River mainstem, Green=Green River mainstem, Gunnison=Gunnison River mainstem, San Juan Trib=tributaries of the San Juan River, San Juan Main=San Juan River mainstem, Monument=Monument River). ................................................................. 21

Figure 2.3. Unrooted neighbor-joining tree of all samples. Colors represent the genotype of the individual: Blue=BHS, Red=FMS, Yellow=WHS, Orange=FMSxWHS, Green=BHSxWHS, Purple=BHSxFMS. Samples are labeled by watershed (COL=Colorado, GR=Green, GUN=Gunnison, SJR=San Juan, SJRTrib=San Juan tributaries, MON=Monument Creek). ................................................................. 22

Figure 2.4. Plots of a principal coordinates analysis of all individuals sampled using the two genetic marker types: AFLPs (a) and the mitochondrial marker ND5 (b) and hybrid index estimations (c). Each point is an individual fish. The amount of variance explained by PCO axes 1 and 2 are 23% and 11%, respectively. Colors represent the genotype of the individual as determined by the Structure analysis: Blue=BHS, Red=FMS, Yellow=WHS, Orange=FMSxWHS, Green=BHSxWHS, Purple=BHSxFMS. Hybrids were considered to be those individuals with an assignment of less than 85% to one cluster (species). In (c), each point is an individual plotted on a scale of 0 to 1, with 0 representing 0% genetic contribution from the species listed on the top of the y-axis and 1 representing 100% contribution coming from the species listed at the top of the y-axis (B=Bluehead, W=White, F=Flannelmouth). Confidence levels calculated by GenoDive are shown for each estimation. Hybrid class assignments given to each individual by NewHybrids is shown on the left, with boxes representing F1 hybrids and diamonds representing F2 hybrids. Points colored in white indicate those that had a posterior probability of <0.90 of being its assigned hybrid class. ....... 24

Figure 2.5. Relative proportions of species and hybrid classes found in each subdrainage based on an analysis using NewHybrids. (B=bluehead, BF=bluehead-flannelmouth hybrid, BW=bluehead-white hybrid, F=flannelmouth, FW=flannelmouth-white hybrid, W=white, Colorado=Colorado...
River mainstem, Green=Green River mainstem, Gunnison=Gunnison River mainstem, San Juan Tributaries=La Plata River, Long Hollow Creek, Mancos River, San Juan=San Juan River mainstem, Monument=Monument River).

Figure 2.6. Relative proportions of species caught in past surveys at the Redlands Fish Ladder on the Gunnison River. Relative proportion was calculated as the percent of the total number of suckers caught in each survey and these were averaged across all surveys in a given year. BHS=Bluehead, FMS=Flannelmouth, WHS=White, BxW=Bluehead x White hybrid)

Figure 3.1. Phylogeny and mouth morphologies of the three sucker species. Phylogeny is based on mitochondrial sequence divergence as described in Doosey et al. (2009). Mouth morphologies are one of the features used to identify these three species in the field. The bluehead sucker has a prominent hard ridge, which the other two species lack. The flannelmouth has a relatively fleshy mouth with enlarged, lower lips compared to the bluehead and white suckers.

Figure 3.2. Clustering and alpha diversity of all samples. Samples are clustered based on principal coordinates analysis (PCoA) of unweighted Unifrac distances (a). Each point represents an individual sample and is colored by sample type. Collection locations are indicated by shapes. A boxplot shows the distribution of bacterial diversity by sample types (b). Diversity was measured as Faith’s Phylogenetic Diversity (PD). *p<0.05 after Bonferroni correction (Kruskal-Wallis test).

Figure 3.3. Main factors affecting the skin bacterial community composition. A heatmap shows the relative abundances of bacterial classes (phylum in parentheses) present in fecal communities across species (a). For each sample type, OTUs with an average relative abundance of greater than 0.1% of total sequences were summarized to the bacterial class level. Taxa found to be significantly different in abundance among blueheads (BHS), flannelmouth (FMS) juveniles (J), flannelmouth adults (A), and white suckers (WHS) are bolded and starred (Kruskal-Wallis test, FDR corrected p<0.05). A boxplot shows the distribution of bacterial diversity in fecal samples by species and age class (b). Diversity was measured as Faith’s Phylogenetic Diversity (PD). Skin communities are clustered based on a PCoA of unweighted (c, main panel) and weighted (c, right panel) UniFrac distance matrices. Each point represents an individual skin sample. Samples are colored by the species identity and symbols indicate the age class of the fish host. The percentage of the variation explained by the plotted principal coordinates is indicated on the axes.

Figure 3.4. Summary of the bacterial composition of each sample type. Genera abundant at a relative abundance of >0.1% of the total dataset are shown. Fecal samples were split into groups according to species and age groups; skin and water samples were divided into groups according to sampling locations. Each bar represents the average proportions across all individuals in the group. The number of individuals in each group is shown above each bar. Classifications indicate the genus of the bacteria followed by their phylum classification in parentheses; for unclassified taxa, family classification is also provided.
Figure 3.5. Main factors affecting the skin bacterial community composition. A heatmap shows the relative abundances of bacterial classes (phylum in parentheses) present in skin communities across sampling locations (a). For each sample type, OTUs with an average relative abundance of greater than 0.1% of total sequences were summarized to the bacterial class level. Taxa found to be significantly different in abundance among the Colorado River (COL), Government Highline (GH), and Redlands (RFL) sampling locations are bolded and starred (Kruskal-Wallis test, FDR corrected p<0.05). A boxplot shows the distribution of bacterial diversity in skin samples by sample location (b). Diversity was measured as Faith’s Phylogenetic Diversity (PD). *p<0.05 after Bonferroni correction (Kruskal-Wallis test). Skin communities are clustered based on a PCoA of unweighted (c, main panel) and weighted (c, right panel) UniFrac distance matrices. Each point represents an individual skin sample. Samples are colored by sampling location and symbols indicate the species identity of the fish host. The percentage of the variation explained by the plotted principal coordinates is indicated on the axes.

Figure 3.6. Comparison of similarities between skin and water samples from the same versus different locations (a-b) and source environment proportions for all skin samples estimated using the 36 fecal samples and each water sample as source environments (c-d). Similarity was assessed as the number of shared OTUs (a) and unweighted UniFrac distances (b). Bar charts show estimates of the mean proportions derived from likely sources for each sample using 100 draws from Gibbs sampling (d) and averaged values for samples from the same location (c). COL=Colorado River open site, GH=Government Highline Fish Ladder, RFL=Redlands Fish Ladder. *p<0.001 after Bonferroni correction (Kruskal-Wallis test).
CHAPTER 1
Introduction

The recent loss and continuing threats to biodiversity has made conservation biology an increasingly important field in recent years. Although ecological, political, and economic considerations are often of primary concern when it comes to wildlife conservation, the goals of conservation research are firmly embedded within the field of evolutionary biology, as they ultimately also require an understanding of the genetic and environmental factors that affect an individual's fitness and in turn the viability of a population. For instance, the genetic characteristics of populations can have profound impacts on their long-term ability to fare with environmental change, invading species and disease (reviewed in Hughes et al. 2008). Thus, processes that affect the genetic composition of a population and the resulting consequences are of great interest to both evolutionary biologists and conservationists.

One powerful mechanism for genetic change is hybridization, in which two genetically distinct populations interbreed and produce viable offspring. The importance of hybridization has been increasingly emphasized in recent decades (Arnold 1992; Rhymer & Simberloff 1996; Allendorf et al. 2001; Mallet 2005) because of the kinds of evolutionary change it can create on a relatively short time scale. Because hybridization can instantly create new combinations of genes, it can rapidly facilitate evolutionary novelty and diversification (Lewontin & Birch 1966; Arnold 2004; Stelkens & Seehausen 2009; Whitney et al. 2010). However, it can also result in genetic homogenization or the extinction of species through a number of mechanisms, with small populations of rare species being particularly at risk (Rhymer & Simberloff 1996). All are reasons the consequences of hybridization has received much attention in conservation research.
While the effects of hybridization have been widely appreciated in conservation biology, the potential effects of an individual’s microbiome (the collection of microbial organisms and their genes) have thus far received considerably less focus in conservation issues. This is surprising because we are quickly learning that host-microbe interactions can have a significant impact on host health and fitness. Microbial commensals play crucial roles in the development, nutrition, and immune capabilities of the host (reviewed in McFall-Ngai et al. 2013). We now also know that this extended host phenotype can have potential fitness effects on an even shorter time frame than hybridization. While an individual’s genotype is a static trait, the microbiome represents a highly alterable phenotype, for example by changes in diet and environmental conditions (Turnbaugh et al. 2009; Spor et al. 2011). Thus, microbiome studies are becoming more applicable and relevant to conversation biology. In fact, an appreciation for this potential was recently noted in the journal Conservation Biology (Redford et al. 2012), wherein the authors speculated that differences in microbiomes may affect interactions between non-native and native species or underlie some of the challenges of managing animals in certain captive breeding and reintroduction programs. Consequently, a better understanding of both of these factors--hybridization and animal-associated microbial communities--is essential, and has significant implications for the conservation of natural systems.

**Hybridization as a conservation problem**

Hybridization is an event that occurs frequently in nature, and likely an important source of genetic variation in natural populations (Arnold 1992). However, it is also often an unintended result of human-mediated introductions of non-native species and has been shown to be a major cause of native species declines (Rhymer & Simberloff 1996). Especially when native species population sizes are low, hybridization can lead to their
effective extinction (at least taxonomically speaking since genes from that species may persist in the existing population). For example, the grey duck population of New Zealand is rapidly declining and likely to become extinct because hybridization with the introduced mallard duck has led to the formation and spread of an increasingly homogenous hybrid population (Rhymer et al. 1994).

Outcomes of hybridization can vary widely among animal systems and even among populations of the same system depending on both the environmental context and the potential genomic interactions between the parent taxa. In some cases, hybridization can introduce maladaptive alleles into a population, in which case individuals are unfit for a particular environment in terms of either survival or reproduction (Arnold et al. 1999). Unfortunately, even if hybridization results in offspring with decreased or equal fitness relative to the parental species, the decline and eventual extinction of native genotypes can result, particularly if introduced species outnumber the natives (Rhymer & Simberloff 1996). However, in other cases, hybridization can infuse adaptive alleles into a population or increase genetic diversity, allowing for a greater range of ecological adaptation (e.g. Whitney et al. 2006; Kays et al. 2010).

In such cases where hybridization represents a path towards a more fit population, difficult conservation issues arise. Should protection be limited to ‘pure’ genotypes or should hybridization instead be considered a way of conserving successful and potentially adaptive gene combinations? In fact, for some endangered species suffering from inbreeding, hybridization with a closely related species or subspecies can be used as a conservation management strategy, and this type of “genetic rescue” has already been highly successful in some cases (e.g. the Florida panther, Johnson et al. 2010). However, for many wildlife systems, information quantifying the effects of introduced species and hybridization on the potential survival of species is necessary before such management decisions are made. In
the first chapter, I provide this gap in information for one system of hybridizing fish species by examining the effects of hybridization with an introduced non-native on two fish species native to the Upper Colorado River Basin.

A need to better understand animal-associated microbial communities

All animals live in close association with a diversity of microorganisms, and biologists have long appreciated the roles that certain microbes play in animal health. However, the advent of new molecular techniques has made the study of entire communities increasingly feasible and accessible, revealing the true diversity, ubiquity, and functional roles of the microorganisms associated with animals. Much of recent microbiome research has focused specifically on the guts of humans and model organisms, providing important insights into the various roles of gut microbiota. For example, variation in gut microbial composition has been implicated in the development of disorders including obesity, diabetes, and inflammation (McKenna et al. 2008; Turnbaugh et al. 2008; Burcelin et al. 2011; Ridaura et al. 2013). However, a growing collection of studies on animal systems are demonstrating the importance of gut commensals on a suite of other functions affecting host fitness including digestion and nutrition, immunity, and development (McFall-Ngai et al. 2013). Some studies even link the gut microbiota to animal behavior (Collins et al. 2012) and speciation processes (Brucker & Bordenstein 2013).

Resident microbiota of other body environments are shown to provide similarly important roles. In humans for instance, one commensal skin bacterium has been shown to be able to selectively inhibit skin pathogens (Cogen et al. 2010). Similarly, certain skin microbiota of fish are thought to help prevent colonization by other microbes including pathogens through competition or a similar mechanism (Westerdahl et al. 1991; Chabrillón et al. 2005). Studies of amphibian skin microbes suggest that susceptibility to disease may
be mediated in part by their skin microbial communities (Woodhams et al. 2007; Harris et al. 2009; Bletz et al. 2013).

Given the various ways that microorganisms can affect host fitness, the applicability of microbiome research to conservation efforts seems logical. However, for many animal systems, particularly those in the wild, there is still a need for basic descriptions of the range of natural diversity and composition of the microbiota associated with their animal hosts as well as a basic understanding of the suite of factors that influence their assembly and flux. In the second chapter, I address this issue by investigating the diversity of microbes associated with the gastrointestinal tracts and skin mucus of the three fish species described below.

Study System

Fish species in particular have been shown to be highly sensitive to the effects of introduced species, which often exacerbates the common problems caused by fragmentation and habitat alteration associated with anthropogenic disturbance (Colautti et al. 2003). Subsequent hybridization with non-native species is a growing concern because hybridization in general is documented to be more common among freshwater teleosts than in other vertebrates (Hubbs 1955; Campton 1987). Heavily affected by both anthropogenic activity and non-native species introductions, the native fish community of the Colorado River Basin (CRB) is considered a highly threatened fish assemblage (Minckley & Deacon 1968). The bluehead sucker (Catostomus discobolus) and flannelmouth sucker (C. latipinnis) are two fish species native to this drainage that have declined dramatically in both distribution and abundance, and only occupy approximately 50% of their historic range (Bezzerides & Bestgen 2002). Although their decline has been attributed in part to
habitat alterations and fragmentation, interactions with non-native fish have also been implicated (Bezzerides & Bestgen 2002).

The white sucker, *C. commersoni*, is one such invading species. Introduced from its native range east of the Continental Divide, this species now occupies several of the streams home to the bluehead and flannelmouth suckers and has been reported to hybridize extensively with both native species (Douglas & Douglas 2007a; McDonald *et al.* 2008). A closely related species to the two natives, the white sucker also shares similar life history and ecological traits to the two native species, including spawning timing and habitat preferences, as well as diet preferences. Currently, several basic questions about the threats of hybridization with non-native species remain unanswered. For example, it is unclear in which conditions and to what extent hybridization will occur when both native species exist in sympatry with the non-native white sucker. Agency managers have acknowledged that an important step in managing populations of these native species will be to reduce or eliminate the threat of hybridization with non-native suckers. Current management efforts include removal of white suckers from stream reaches. However, given the widespread distribution and abundance of the white sucker, this will not be easy or perhaps even feasible, particularly in the larger sub-drainages of the Upper CRB. Thus, information on where and when hybridization with the white sucker is likely to have the most severe consequences would be useful in prioritizing areas of native species conservation and management.

We currently lack fundamental information regarding general patterns of hybridization among these three species, and the factors affecting their associated microbiota, each with the ability to affect the fitness of the species. Therefore, my objectives were to: 1) quantify the distribution of hybrid types across a large geographic portion of the
Upper Colorado River Basin, 2) determine whether these patterns were related to estimates of relative white sucker population size as determined by 20 years of survey data, 3) describe the diversity and composition of the gut and skin microbial communities, and lastly, using this information, 4) identify the factors correlated with patterns of microbial diversity.
CHAPTER 2

An analysis of patterns of hybridization between the introduced white sucker (*Catostomus commersoni*) and native bluehead (*C. discobolus*) and flannelmouth (*C. latipinnis*) suckers in the Upper Colorado River Basin

INTRODUCTION

Native fish populations of North America have experienced rapid declines in the last century (Moyle & Leidy 1992). A study of extinctions during a 100-year period between 1889 and 1989 indicated that an estimated 40 fish taxa were lost (Miller *et al.* 1989). Although several factors contribute to declines, nonnative fish species have been implicated in nearly 50% of endangered species listings, second only to habitat loss (Magnuson *et al.* 1998). In fact, human-mediated introductions of nonnative species are considered to be one of the biggest threats to native biodiversity (Vitousek *et al.* 1997; Simberloff 2003; Torchin *et al.* 2003). Once species are introduced into new environments, consequent contact with historically isolated but closely related native species can often result in hybridization, since reproductive barriers between allopatrically evolved species tend to be weaker than between sympatrically evolved species (Coyne & Orr 2004). Hybridization of introduced with native species can thus pose a serious threat to the persistence of native species and has been found to be a factor in at least 38% of the recorded extinctions of North American fish species (Miller *et al.* 1989).

Studies suggest that rates of hybridization will increase as a result of accelerating anthropogenic change (e.g. Seehausen *et al.* 2008); hence, the consequences of hybridization are important to consider in the conservation of native species (Rhymer & Simberloff 1996). However, understanding the circumstances in which introduction can lead to introgressive hybridization (interspecific mating where genes are transferred between species as a result
of backcrossing events) between introduced and native taxa is challenging because the outcomes can vary depending on environmental factors and the potential genomic interactions between the parent taxa. When hybridization does occur, the extent often depends on the amount of reproductive isolation between the hybridizing populations, with potential outcomes including the loss of local adaptations, the generation of novel adaptations, the extinction of species, or in the right conditions, the formation of new hybrid species (Lewontin & Birch 1966; Allendorf et al. 2001; Burke & Arnold 2001; Rieseberg et al. 2003; Stelkens & Seehausen 2009). Even low levels of hybridization can introduce locally adaptive alleles into invading populations, facilitating their advancement into new environments. For these reasons, understanding how and why patterns of hybridization may vary across populations and locations has important implications for the conservation of declining species and may help predict their trajectories.

Compared to other vertebrates, hybridization and introgression are relatively more common among freshwater teleost fish (Hubbs 1955; Campton 1987). In North America, fishes in the teleost family Catostomidae (commonly known as suckers) exhibit both high degrees of diversity and geographic endemism (Smith 1966; Lee et al. 1980). Most of the 23 currently recognized species in the genus Catostomus are endemic to areas of the western United States (Smith 1992). However, hybridization commonly occurs where historically separated species come into contact (Smith 1966).

The bluehead sucker *Catostomus discobolus* (BHS), and the flannelmouth sucker, *Catostomus latipinnis* (FMS) are two species of catostomids native to the Colorado River Basin (CRB), west of the Continental Divide. Their ranges overlap extensively, with the two occurring sympatrically in several drainages in the Upper CRB. Reports indicate that BHS and FMS populations were historically common to abundant in all or parts of the CRB (Sigler & Miller 1963; Jordan & Evermann 1969; Minckley 1973) However, evidence
indicates that these populations are now declining in both number and distribution. As of 2002, they were estimated to only occupy approximately 50% of their historical ranges (Bezzerides & Bestgen 2002). Due to the reported declines of these species, a range-wide conservation agreement was collectively developed in 2004 by six western states and multiple government agencies, with the main goal of ensuring their persistence (Colorado River Fish and Wildlife Council 2006). Although factors such as habitat alteration and population fragmentation due to dam construction and operation have also been implicated, their decline has been determined to be due in part to hybridization with non-native fishes (Martinez et al. 1994; Bestgen & Crist 2000; Bestgen et al. 2006). One such species is the non-native white sucker, *Catostomus commersoni* (WHS). Introduced into the CRB from their native range east of the Continental Divide, white suckers now occupy almost all of the range of BHS and FMS suckers. Although the white sucker did not become widespread in the CRB until the late 20th century, hybridization between the native suckers and WHS likely began much earlier, when juveniles were inadvertently introduced during trout stocking and as fishing bait (Vanicek 1970; Tyus et al. 1982; Woodling 1985). Evidence suggests that multiple introductions of the white sucker have occurred over several decades (McPhee & Turner 2009), allowing for its rapid expansion across the CRB (Roman & Darling 2007). Moreover, partial overlap in spawning times, spawning behavior, and spawning habitat preferences among the three species provide some opportunities for hybridization. Given these circumstances, it is not surprising that the introduced white sucker has been reported to hybridize readily with both native species in areas of sympatry (Douglas & Douglas 2007a).

One study has even shown that the white sucker can not only hybridize with both native catostomid species, but may also be facilitating gene flow between the two native species by acting as a genetic bridge (McDonald et al. 2008). Out of 161 fish sampled,
McDonald and colleagues identified individuals with genetic contributions from all three species and inferred that these hybrids resulted from three-way hybridization (e.g. an F1 hybrid between two species crossed with the third species). This finding was intriguing because although bluehead x flannelmouth hybrids are occasionally found, with an estimated 13% sequence divergence based on mitochondrial DNA (McDonald et al. 2008; Doosey et al. 2010), introgression of genes between these two native species is not likely. In fact, for many species with this amount of genetic divergence, successful hybridization is no longer possible (Coyne & Orr 2004).

The existence of three-way hybrids is greatly disconcerting and has significant conservation implications for the native species. The formation of three-way hybrids suggests that hybridization may pose a greater threat to native populations than previously thought by providing an additional route for genetic degradation of the native species. Increased invasive ability is also a possibility because these hybrids have the potential to carry an even greater range of novel genotypes than two-species hybrids. However, it is important to note that the study conducted by McDonald et al. was restricted to one of the many drainages of the Upper CRB where the three species co-occur, and whether this is a general pattern that occurs in all areas of sympatry is unknown.

Thus, our aim was to examine the patterns of hybridization and introgression among these species across a broader geographic range of the native species. We were particularly interested in determining the distribution of each of the hybrid types between species pairs and among watersheds. Fish of pure parent species and their putative hybrids were collected from 11 locations across the Upper CRB and genotyped. We used admixture analyses to identify the genetic compositions of putative hybrids, which allowed us to determine the extent of hybridization of the native species with the introduced species and revealed information regarding patterns of reproductive isolation and mating in recent
generations. Because the strength of reproductive isolation between species is generally expected to increase with phylogenetic distance, we hypothesized that the extent of introgression among parent species would vary inversely with distance among species pairs. Furthermore, we obtained survey records from Colorado Parks and Wildlife and the United States Fish and Wildlife Service from the Redlands Fish Ladder, a structure near the confluence of the Gunnison and Colorado Rivers in Grand Junction, Colorado, which we used to assess white sucker relative abundance over the last 15 years. We use these data to assess whether certain hybridization patterns can be explained by the history of white sucker abundance relative to the native suckers. Results of this work will be important for identifying areas where hybridization is most prevalent and establishing management and conservation strategies to benefit the two native catostomid species.

MATERIALS AND METHODS

Sampling and prior sampling data

Eleven locations were sampled covering a large portion of the range of the two native species within the Upper Colorado River Basin (UCRB). These included sites along the Gunnison and Colorado Rivers in Colorado, the Green River in Utah, and several streams within the San Juan River Basin. Additionally, we included samples of white suckers from one population east of the Continental Divide where it is native (Monument Creek, Colorado Springs, CO) (Fig. 2.1; Table 2.1). In the San Juan River Basin, the white sucker is present in the San Juan River mainstem although in low numbers, and absent in some, although not all, of the smaller tributaries. Both site types within this basin were sampled. All three species occur sympatrically throughout the other subdrainages of the UCRB. Between the years of 2005 and 2009, a total of 397 tissue samples were collected from bluehead, flannelmouth, and white suckers, as well as putative hybrids. For this
study, sampling was not performed randomly but rather focused on putative hybrids. Sample sizes thus do not reflect relative abundance of the various species and hybrids. Samples were collected under Colorado scientific permit #08AQ1026, University of Colorado protocol #1103.05, and the permits of Colorado Parks and Wildlife and Utah Division of Wildlife Resources.

Figure 2.1. Sampling locations in the Upper Colorado River Basin and the species found at those sites. One location outside of the UCRB (Monument Creek in Colorado Springs, CO, represented in yellow) was sampled. Numbers correspond to the locations listed in Table 2.1. The green star indicates the location of the Redlands Fish Ladder, a fish passage structure located on the Gunnison River near its confluence with the Colorado River. This location was not sampled, but survey data from this locality were used.
Table 2.1. Sampling locations and sample sizes by species. Numbers correspond to the locations shown in Figure 2.1. BHS=Bluehead, FMS=Flannelmouth, WHS=White, BxW=Bluehead x White hybrids, FxW=Flannelmouth x White hybrids, BxF=Bluehead x Flannelmouth hybrids.

<table>
<thead>
<tr>
<th>Location</th>
<th>Main River</th>
<th>BHS</th>
<th>FMS</th>
<th>WHS</th>
<th>BxW</th>
<th>FxW</th>
<th>BxF</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Colorado River near Clifton (CR)</td>
<td>Colorado River</td>
<td>15</td>
<td>22</td>
<td>9</td>
<td>16</td>
<td>8</td>
<td>2</td>
<td>72</td>
</tr>
<tr>
<td>2. Colorado River near Grand Junction (CRGJ)</td>
<td>Colorado River</td>
<td>15</td>
<td>15</td>
<td>12</td>
<td>12</td>
<td>3</td>
<td>0</td>
<td>57</td>
</tr>
<tr>
<td>3. Green River near Jensen (GR)</td>
<td>Green River</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>4. Green River near Desolation Canyon (GRDC)</td>
<td>Green River</td>
<td>12</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>5. Green River near Split Mountain (GRSM)</td>
<td>Green River</td>
<td>6</td>
<td>6</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>6. Gunnison River between Dominguez and Bridgeport (GUN)</td>
<td>Gunnison River</td>
<td>9</td>
<td>12</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>34</td>
</tr>
<tr>
<td>7. Nenahnezad Fish Passage (NFP)</td>
<td>San Juan River</td>
<td>28</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>58</td>
</tr>
<tr>
<td>8. Mancos River (MAN)</td>
<td>San Juan River</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>9. Long Hollow Creek (LHC)</td>
<td>San Juan River</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>10. La Plata River (LPR)</td>
<td>San Juan River</td>
<td>1</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>11. San Juan River near Farmington (SJR)</td>
<td>San Juan River</td>
<td>20</td>
<td>18</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>43</td>
</tr>
<tr>
<td>12. Monument Creek (MON)</td>
<td>Arkansas River</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>122</td>
<td>159</td>
<td>60</td>
<td>34</td>
<td>18</td>
<td>4</td>
<td>397</td>
</tr>
</tbody>
</table>

Survey data kept by the United States Fish and Wildlife Service were provided by Colorado Parks and Wildlife and used to estimate the relative abundance of the introduced white sucker between the years of 1996 and 2012. These data included information from 3,849 surveys with records of the number and species identification of fish caught during each survey at the Redlands Fish Ladder, a fish passage structure located on the Gunnison River near its confluence with the Colorado River (Figure 2.1). Survey data for this location
was chosen because the site is located near sample sites and is a location where surveys had been consistently conducted numerous times over multiple years using the same method. Biases introduced by variation in methodology or sampling frequency should therefore have been minimized and we are confident that the true relative abundances of fish species are reflected by these data.

In order to summarize these data into more manageable and meaningful units, we first excluded all surveys that resulted in the catch of less than 10 total fish. The data were further filtered to only include the three species of interest (the bluehead, flannelmouth, and white suckers, along with their hybrids). Species counts were then standardized and summarized by calculating the relative proportions of each species caught for each survey. For each species, these proportions were then averaged across all surveys for every year.

**DNA extraction and genotyping**

Total DNA was extracted from 2mm x 2mm pieces of collected fin tissue using the Qiagen DNAeasy tissue kit. DNA concentrations were quantified using a NanoDrop 2000 (ThermoScientific). Individuals were identified in the field as BHS, FMS, WHS, or hybrids through an analysis of morphological traits. Identities were either confirmed or corrected by genotyping, using AFLP (amplified fragment length polymorphism) and mitochondrial DNA (mtDNA) markers. The AFLP method first developed by Vos et al. (1995) allows for the development of hundreds of genetic markers across the genome without prior sequence information, which makes it particularly useful for the study of non-model organisms (Meudt & Clarke 2007). Due to its widespread applicability, it has been widely utilized for studies of population structure and the identification of hybrid individuals. Mitochondrial DNA (mtDNA) on the other hand has limited applicability for hybrid detection because it is only inherited maternally. However, mtDNA markers, when combined with nuclear
markers such as AFLPs, can be used to infer directionality of mating (i.e. which genders of parental species were involved in the hybridization event). We therefore used a combination of AFLP genotype data and mitochondrial sequence data.

The mitochondrial gene encoding for NADH dehydrogenase subunit 5 (ND5) was amplified and sequenced using the primers 5’-TCAGCTTCAAGTTCGACCAC and 5’-AAACCAGGCTTGCTCTACGA. PCR was carried out using 1μL of DNA, 10x PCR buffer, 1mM MgCl2, 0.2mM dNTPs, 1 pmol of primers, and 1 unit of Taq DNA polymerase (Promega Co.). PCR conditions were as follows: 35 cycles of 94°C for 45 s, 56°C for 30 s and 72°C for 90 s, followed by a final 4 min extension at 72°C. Prior to sequencing, PCR products were cleaned using 1 unit of exonuclease and 1 unit of shrimp alkaline phosphatase (New England Biolabs) per 10 μl of product through incubation at 37°C for 15 minutes and then at 80°C for 15 min to deactivate the enzymes. Amplicons were sequenced by Functional Biosciences, Inc. (Madison, WI) from both the forward and reverse directions on an ABI 3730. Sequences were assembled and aligned using the program Sequencher 4.7 (Gene Codes).

AFLP markers were generated and amplified using a modified version of the Vos et al. (1995) protocol implemented by McDonald et al. (2008). Three selective combinations of primers were used: EcoRI-ACT / MseI-CTA, EcoRI-AAC / MseI-CTT, and EcoRI-AAT / MseI-CAT. Fragment analysis was performed by the Nevada Genomics Center in Reno on an ABI 3130 capillary sequencer, and AFLP fragments were binned and scored using Genemapper v4.0 (Applied Biosystems Inc.). Fragments of sizes between 60 and 400 bp were scored as present (1) or absent (0) at each locus for each individual. To reduce the rate of errors, all samples were run at the same time and scored twice. Bands that were not scored in both passes were excluded from downstream analyses.
Species identification and hybrid assignment

Once AFLP data were generated, we used the program AFLP-SURV (Vekemans et al. 2002) to estimate expected heterozygosity and overall amounts of population differentiation ($F_{ST}$) among the three species. The program STRUCTURE v2.2 (Pritchard et al. 2000; Falush et al. 2007) was used to infer the total number of genotypic clusters (K) and to estimate probabilities of population assignment (Q) and genetic admixture for each individual. We applied an admixture model that assumes no prior population information, gene flow among populations, and correlated allele frequencies. Five independent simulations were run for each value of K from 2 to 10. Changes in the log probability (L(K)) between K values were used to find the optimal number of clusters as suggested in Evanno et al. (2005). We then reran the analysis assuming this number of genetic clusters. Hybrids were considered to be those individuals with an assignment of less than 85% to one cluster (species).

We also used the program GenoDive (Meirmans & Van Tienderen 2004) to verify the hybrid composition of individuals detected by Structure and the program NewHybrids (Anderson 2008) to infer the hybrid class of these individuals as F1, F2, or backcrossed/later generation hybrids. GenoDive uses the maximum likelihood approach of Buerkle (2005) to quantify the genetic contribution of parental species to putative hybrids (i.e. hybrid index). Utilizing a Bayesian probabilistic approach, NewHybrids uses allele frequencies in the parent populations to estimate the probability that the hybrid individual is of a certain hybrid category. For each pair of species, the individuals from the known pure populations (Sites #7, 8, 9, 10, and 12) were assigned pure genotype classes. Calculations were run using Jeffreys-like priors for estimating allele frequencies and mixing proportions with a burn-in of 5000 iterations followed by 50000 sweeps for sampling from the posterior distribution.
AFLP data were used to create a matrix of pairwise genetic distances among individuals using the RestDist package in PHYLIP v3.6 (Felsenstein 2005), which implements a modified approach of Nei & Li (1979). Using this distance matrix, we performed a principal coordinates analysis (PCoA) and visualized the data using the first two axes. We also created an unrooted neighbor-joining tree with PHYLIP’s Neighbor package and visualized the tree using Topiary Explorer (Pirrung et al. 2011).

Detection of outlier loci

Although most loci are expected to be influenced similarly by genetic drift and gene flow, those under divergent selection (or linked to loci under selection) are expected to display a relatively high level of genetic differentiation (Luikart et al. 2003; Storz 2005; Nosil et al. 2009) and are referred to as ‘outlier loci’. We used the program BayeScan v2.1 (Foll & Gaggiotti 2008) to detect outlier loci that may be under divergent selection between species. Briefly, the method uses differences in allele frequencies between populations to calculate $F_{ST}$ and a hierarchical Bayesian approach to estimate for each locus the posterior probabilities of a neutral model versus one that accounts for selection. For outlier locus detection, we required a posterior probability threshold of >0.99 ($\log_{10}PO > 2$) and a false discovery rate (FDR) of <0.05 % for the model including the effects of selection. BayeScan was run with 20 pilot runs, a burn-in of 10,000 followed by 10,000 iterations each, a sample size of 5000, and a thinning interval of 50. We ran this analysis separately for each of the species pairs using only individuals identified as ‘pure’ parental genotypes.

RESULTS

Genetic differentiation among and within species
We successfully genotyped 375 of the 397 fish representing 11 populations within the Colorado River Basin, plus one additional reference population of the white sucker. AFLP genotyping produced 453 loci, of which 384 (84.8%) were polymorphic. Interestingly, we found relatively high amounts of genetic diversity in the introduced white sucker populations (as assessed by the proportion of polymorphic AFLP loci) despite low sample size (Table 2.2). Although introduced populations are often predicted to have low genetic diversity as a result of genetic bottlenecks and small population size (Roman & Darling 2007), these results seem to be more consistent with the idea that there were indeed multiple introductions of the white sucker to the Upper CRB (McPhee & Turner 2009).
Table 2.2. Summary of the genetic diversity found in each population sampled listed by species (Species) and location (Pop). Genetic diversity measures are provided as the percentage of AFLP loci that were polymorphic (%poly loci) as well as expected heterozygosity (Hj), which is also provided with standard error (s.e.(Hj)) and variance (var(Hj)). See Table 2.1 for abbreviations of species and localities.

<table>
<thead>
<tr>
<th>Species</th>
<th>Pop</th>
<th>n</th>
<th>#poly loci</th>
<th>%poly loci</th>
<th>Hj</th>
<th>s.e.(Hj)</th>
<th>var(Hj)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BHS</td>
<td>CRCL</td>
<td>14</td>
<td>246</td>
<td>54.3</td>
<td>0.1969</td>
<td>0.00868</td>
<td>0.000075</td>
</tr>
<tr>
<td>BHS</td>
<td>CRGJ</td>
<td>14</td>
<td>246</td>
<td>54.3</td>
<td>0.20847</td>
<td>0.00926</td>
<td>0.000086</td>
</tr>
<tr>
<td>BHS</td>
<td>GR</td>
<td>6</td>
<td>223</td>
<td>49.2</td>
<td>0.19011</td>
<td>0.00905</td>
<td>0.000082</td>
</tr>
<tr>
<td>BHS</td>
<td>GRDC</td>
<td>11</td>
<td>218</td>
<td>48.1</td>
<td>0.19134</td>
<td>0.00925</td>
<td>0.000086</td>
</tr>
<tr>
<td>BHS</td>
<td>GRSM</td>
<td>5</td>
<td>240</td>
<td>53</td>
<td>0.21768</td>
<td>0.00952</td>
<td>0.000091</td>
</tr>
<tr>
<td>BHS</td>
<td>GUN</td>
<td>6</td>
<td>299</td>
<td>66</td>
<td>0.20999</td>
<td>0.00951</td>
<td>0.00009</td>
</tr>
<tr>
<td>BHS</td>
<td>LHC</td>
<td>10</td>
<td>250</td>
<td>55.2</td>
<td>0.19119</td>
<td>0.00888</td>
<td>0.000079</td>
</tr>
<tr>
<td>BHS</td>
<td>NFP</td>
<td>26</td>
<td>241</td>
<td>53.2</td>
<td>0.18589</td>
<td>0.00883</td>
<td>0.000078</td>
</tr>
<tr>
<td>BHS</td>
<td>SJR</td>
<td>9</td>
<td>242</td>
<td>53.4</td>
<td>0.17608</td>
<td>0.00856</td>
<td>0.000073</td>
</tr>
<tr>
<td>FMS</td>
<td>CRCL</td>
<td>21</td>
<td>280</td>
<td>61.8</td>
<td>0.21565</td>
<td>0.00901</td>
<td>0.000081</td>
</tr>
<tr>
<td>FMS</td>
<td>CRGJ</td>
<td>13</td>
<td>264</td>
<td>58.3</td>
<td>0.19759</td>
<td>0.00862</td>
<td>0.000074</td>
</tr>
<tr>
<td>FMS</td>
<td>GR</td>
<td>7</td>
<td>253</td>
<td>55.8</td>
<td>0.21475</td>
<td>0.00919</td>
<td>0.000084</td>
</tr>
<tr>
<td>FMS</td>
<td>GRDC</td>
<td>14</td>
<td>260</td>
<td>57.4</td>
<td>0.20163</td>
<td>0.00876</td>
<td>0.000077</td>
</tr>
<tr>
<td>FMS</td>
<td>GRSM</td>
<td>6</td>
<td>254</td>
<td>56.1</td>
<td>0.21398</td>
<td>0.00903</td>
<td>0.000082</td>
</tr>
<tr>
<td>FMS</td>
<td>GUN</td>
<td>12</td>
<td>227</td>
<td>50.1</td>
<td>0.17412</td>
<td>0.00861</td>
<td>0.000074</td>
</tr>
<tr>
<td>FMS</td>
<td>LHC</td>
<td>2</td>
<td>190</td>
<td>41.9</td>
<td>0.19454</td>
<td>0.00893</td>
<td>0.00008</td>
</tr>
<tr>
<td>FMS</td>
<td>LPR</td>
<td>23</td>
<td>248</td>
<td>54.7</td>
<td>0.18233</td>
<td>0.00857</td>
<td>0.000074</td>
</tr>
<tr>
<td>FMS</td>
<td>MAN</td>
<td>8</td>
<td>244</td>
<td>53.9</td>
<td>0.19865</td>
<td>0.00914</td>
<td>0.000084</td>
</tr>
<tr>
<td>FMS</td>
<td>NFP</td>
<td>30</td>
<td>259</td>
<td>57.2</td>
<td>0.17712</td>
<td>0.00856</td>
<td>0.000073</td>
</tr>
<tr>
<td>WHS</td>
<td>CRCL</td>
<td>9</td>
<td>259</td>
<td>57.2</td>
<td>0.20916</td>
<td>0.00898</td>
<td>0.000081</td>
</tr>
<tr>
<td>WHS</td>
<td>CRGJ</td>
<td>12</td>
<td>278</td>
<td>61.4</td>
<td>0.21621</td>
<td>0.0091</td>
<td>0.000083</td>
</tr>
<tr>
<td>WHS</td>
<td>GR</td>
<td>5</td>
<td>248</td>
<td>54.7</td>
<td>0.22774</td>
<td>0.00933</td>
<td>0.000087</td>
</tr>
<tr>
<td>WHS</td>
<td>GRSM</td>
<td>8</td>
<td>252</td>
<td>55.6</td>
<td>0.20762</td>
<td>0.00913</td>
<td>0.000083</td>
</tr>
<tr>
<td>WHS</td>
<td>GUN</td>
<td>9</td>
<td>268</td>
<td>59.2</td>
<td>0.21351</td>
<td>0.00899</td>
<td>0.000081</td>
</tr>
<tr>
<td>WHS</td>
<td>MON</td>
<td>14</td>
<td>240</td>
<td>53</td>
<td>0.21141</td>
<td>0.00938</td>
<td>0.000088</td>
</tr>
</tbody>
</table>

Analyses using Structure indicated that these AFLP markers are robust in their ability to distinctly identify the three species. Individuals from reference populations assigned clearly into three separate groups: white suckers from a population in its native range (Monument Creek) and blueheads and flannelmouths from the tributaries of the San Juan River where whites are absent (Fig. 2.2). However, within species, there appeared to
no subclustering by sample location (Fig. 2.3). A PCoA of the AFLP data also validated species identification and showed a distinct clustering pattern associated with species delineation, with each of the three species in their respective corners and hybrids falling along the axes between the clusters (Fig. 2.4a). The PCoA was also able to accurately reflect phylogenetic distance among the species; the more distantly related native species separated along the first axis, PC1, which captured 23% of the total variance, while the introduced white sucker separated from the two native species along PC2, which captured 11% of the total variance.

Figure 2.2. Genetic cluster assignments from an analysis with Structure of 397 individuals into three groups (K=3) based on 457 AFLP loci. Each vertical bar on the graph corresponds to an individual and the different colors represent the proportional membership in each of the three clusters (Blue=BHS, Red=FMS, Yellow=WHS). Individuals are grouped by sampling locations (Colorado=Colorado River mainstem, Green=Green River mainstem, Gunnison=Gunnison River mainstem, San Juan Trib=tributaries of the San Juan River, San Juan Main=San Juan River mainstem, Monument=Monument River).
Figure 2.3. Unrooted neighbor-joining tree of all samples. Colors represent the genotype of the individual: Blue=BHS, Red=FMS, Yellow=WHS, Orange=FMSxWHS, Green=BHSxWHS, Purple=BHSxFMS. Samples are labeled by watershed (COL=Colorado, GR=Green, GUN=Gunnison, SJR=San Juan, SJR Trib=San Juan tributaries, MON=Monument Creek).
Figure 2.4. Plots of a principal coordinates analysis of all individuals sampled using the two genetic marker types: AFLPs (a) and the mitochondrial marker ND5 (b) and hybrid index estimations (c). Each point is an individual fish. The amount of variance explained by PCO axes 1 and 2 are 23% and 11%, respectively. Colors represent the genotype of the individual as determined by the Structure analysis: Blue=BHS, Red=FMS, Yellow=WHS, Orange=FMSxWHS, Green=BHSxWHS, Purple=BHSxFMS. Hybrids were considered to be those individuals with an assignment of less than 85% to one cluster (species). In (c), each point is an individual plotted on a scale of 0 to 1, with 0 representing 0% genetic contribution from the species listed on the top of the y-axis and 1 representing 100% contribution coming from the species listed at the top of the y-axis (B=Bluehead, W=White, F=Flannelmouth). Confidence levels calculated by GenoDive are shown for each estimation. Hybrid class assignments given to each individual by NewHybrids is shown on the left, with boxes representing F1 hybrids and diamonds representing F2 hybrids. Points colored in white indicate those that had a posterior probability of <0.90 of being its assigned hybrid class.

Patterns of hybridization among species pairs

Where the three species co-occur (Colorado, Green, Gunnison river drainages, and the San Juan River mainstem), hybrids between both native species and the introduced white (BHS x WHS, FMS x WHS), as well as hybrids between the two native species (BHS x FMS), were identified using Structure (Fig. 2.2). In PCoA space, these putative hybrids lay at an intermediate position between the clusters of the parent species. GenoDive and NewHybrids provided results congruent with that of Structure and the PCoA analysis; that is, all approaches identified the same individuals as having admixed genetic ancestry. The greatest number of hybrids were identified between the native bluehead and introduced white (n=34), with fewer between the flannelmouth x white (n=18), and the least between the two native species (n=4) (Fig 2.4a,c). Using NewHybrids, we were able to further categorize these individuals as early generation (F1 or F2), or later generation (backcrossed) hybrids (Fig. 2.4c).

We found that despite the abundance of bluehead x white hybrids detected, these hybrids constitute a relatively narrow range of possible admixed compositions; a majority of
the hybrid individuals (31 of the 34) were classified as early generation hybrids (F1 or F2), only three individuals were identified as backcrosses to the bluehead, and no backcrosses to the white sucker were detected. Furthermore, of the early generation hybrids, F2 hybrids seemed to be much higher in abundance than F1 hybrids. Although there was some uncertainty around the assignments for several of these individuals (posterior probability <0.90, Fig. 2.4c), F2 hybrids were still found to be in greater relative abundance than F1s even after accounting for these uncertainties.

In contrast to the narrow window of hybridization space represented by bluehead x white hybrids, flannelmouth x white hybrids represented a much wider range of genetically admixed combinations in that both early and later generation hybrids were detected. Specifically, backcrossed individuals in each direction were present as well as several F2's, although no individuals were classified as first generation hybrids. Lastly, hybrids between the two native species were rare, with all 4 bluehead x flannelmouth hybrids classifying as F2 hybrids.

Thus, although the most hybrids were found between blueheads and whites, the range of hybrids was widest between flannelmouths and whites, which can be visualized through the results of the PCoA and the hybrid index (Fig. 2.4a,c). In the PCoA plot, because we can consider the axes connecting species clusters as hybrid space, the extent to which this space is filled by individuals gives an indication of the range of hybrids. For instance, flannelmouth x white individuals fell along the full axis between the cluster of white suckers and the cluster of flannelmouth suckers, suggesting that nearly the entire range of hybrids between these two species were sampled. In contrast, the axis between the white and bluehead clusters was only partially filled, suggesting that certain hybrid classes were either absent or present but not sampled.
When this same PCoA is visualized by the mtDNA haplotype of the individuals (Fig. 4b), it is also clear that a majority of the hybrids between the introduced and the two native suckers had the mitochondrial DNA of the introduced sucker (32 of 34 bluehead x white hybrid individuals and 13 of 18 of flannelmouth x white hybrids).

**Variation in hybridization and white sucker occurrence among streams**

Because sampling effort was focused on obtaining putative hybrids, our methods precluded us from being able to measure frequency of hybridization across our sample sites. However, we were able to describe differences in the types of hybrids found at each of these sites. Of all the streams sampled, the Colorado River contained the widest range of hybrids, which included all hybrid categories except FMS-WHS hybrids backcrossed to WHS (Fig. 2.5). The Gunnison River contained very few hybrids, but included both BHS-WHS and FMS-WHS crosses, whereas the Green River contained only FMS-WHS hybrids (n=4). In the San Juan River, a few BHS-WHS hybrids (n=3) were sampled along with one BHS-FMS hybrid.

**Relative abundance of species**

Survey data taken from the Redlands Fish Ladder between 1996 and 2012 show that the relative abundance of the white sucker compared to the native suckers remained low at <10% through the late 1990’s, but increased to much greater numbers in subsequent years (Fig. 2.6). Native sucker abundances on the other hand, remained high through all the surveyed years with blueheads being more abundant than the flannelmouth in most years. Bluehead x white hybrids were the only hybrid type reported at this location, with abundances persisting at low levels (<10%) throughout the surveyed period. However, their
abundance appears to follow a similar trend to white sucker abundances, showing a notable increase around the year 2000.

**Outlier loci**

Ten loci were detected to be under selection between the flannelmouth and white suckers compared to 7 loci between the bluehead and white. Three of these loci were shared between the two comparisons. No loci were detected to be under divergent selection between the two native suckers, the bluehead and flannelmouth.

![Figure 2.5. Relative proportions of species and hybrid classes found in each subdrainage based on an analysis using NewHybrids. (B=bluehead, BF=bluehead-flannelmouth hybrid, BW=bluehead-white hybrid, F=flannelmouth, FW=flannelmouth-white hybrid, W=white, Colorado=Colorado River mainstem, Green=Green River mainstem, Gunnison=Gunnison River mainstem, San Juan Tributaries=La Plata River, Long Hollow Creek, Mancos River, San Juan=San Juan River mainstem, Monument=Monument River).](image-url)
DISCUSSION

In this catostomid system, overlap in preferences for spawning habitat, behavior, and spawning times provide some opportunities for interspecific hybridization. An absence of other reproductive barriers such as selection against hybrids could thus facilitate the formation of a hybrid swarm. However, our findings indicate that the potential for genetic exchange between species depends on the pair considered and may be limited depending on their phylogenetic distance (as a proxy for the amount of reproductive isolation between species). Moreover, we found a corresponding lack of certain hybrid classes, suggesting that some hybrid combinations may be inviable, selected against, or unfit. For example, we found the largest gap in hybrid types between the two native species, the bluehead and...
flannelmouth. The absence of later generation hybrids between these species suggests that hybridization does not extend beyond the F2 generation. One possibility is that F2 hybrids may be sterile, a trait often exhibited by F2 progeny due to the expression of incompatible gene combinations (a.k.a. Muller-Dobzhansky model) (e.g. see Endler 1977). Thus, if early generation hybrids are indeed an evolutionary dead end for genes, it seems that gene flow is highly restricted between the bluehead and flannelmouth sucker species. Moreover, we can interpret the excess of F2 hybrids between these native suckers as an indication that there may be preferential mating among F1 progeny, as opposed to mating with individuals of either parental species. However, it is equally likely that backcrossing may result in inviable or highly unfit offspring. Regardless of the cause, such patterns of hybridization suggest that these two native species are strongly reproductively isolated, likely as a result of evolving in sympatry (Coyne & Orr 2004).

The potential for gene flow between each of the native species with the introduced white sucker, on the other hand, appears to be much greater, particularly for the most closely related species pair, the native flannelmouth and introduced white. In addition to early generation hybrids, we found what appeared to be a gradient of hybrid individuals, which included a range of later generation hybrids (Fig. 2.4a,c). Thus, it appears that continued hybridization between the white sucker and the flannelmouth sucker may indeed eventually lead to a hybrid swarm. These results are largely consistent with the findings of previous studies (Douglas & Douglas 2007a; McDonald et al. 2008). However, our results also show that there is variation in the types of hybrids that form and survive among drainages (Fig. 2.5).
Differing patterns of hybridization among streams

Hybrid distributions differed among sites, with the Colorado River containing the widest range of hybrid classes, followed by the Green River, and the Gunnison River sites. Perhaps most strikingly, individuals with genetic contributions from all three species were absent in these sampled streams, a compelling contrast to what was found in the Muddy Creek drainage in Wyoming by McDonald et al. (2008). Whereas several individuals assigning to all three species were detected in the Muddy Creek (n=3 out of 137 individuals sampled from areas of co-occurrence), no such hybrids were found in the streams sampled in this study (n=0 out of 277 individuals sampled from areas of co-occurrence). Thus, it appears that the potential for the introduced white sucker to serve as a bridge species between the two native species may not generally be applicable to all areas of co-occurrence, but is rather a feature specific to the Muddy Creek drainage.

One alternative explanation is that these differences in hybrid distribution among streams are simply a function of differences in sampling effort, and a proportions test comparing our study to that of McDonald et al. (3/137 versus 0/277; p>0.05) supports this explanation. Although differences in hybrid distribution appear to be present among subdrainages, we do not currently know whether more intense sampling will reveal that all hybrid types are actually present in all systems where the three species co-occur. However, this thesis is not the only study to suggest that there may be geographic variation in sucker hybridization in this system. At least one stream in Wyoming (Henry’s Fork) is reported to have no or few hybrids despite supporting populations of all three species (Gelwicks et al. 2009). Given that our sampling effort focused specifically on putative hybrids, more plausible causes of geographic variation in hybrid distributions may be differences in genetic variation or demographic properties among populations, environmental conditions among streams, or a combination of these factors.
Across our sample sites (which we note did not include the Muddy Creek), we found weak geographic structure within species. This is consistent with previous studies, which reported genetic substructuring on a large geographic scale (i.e. across the entire CRB), but found that populations in large portions of the Upper Colorado River Basin exhibited evidence of high levels of gene flow and thus could be treated as one management unit (Douglas & Douglas 2007b, 2007c; Hopken et al. 2013). Hence, variation in hybrid distributions is likely not a function of population-specific interactions between genetic backgrounds, but rather differences in other factors among locations such as introduction history of the white sucker, relative population sizes of the species, or environmental conditions.

Studies of hybridization between native and introduced salmonids have demonstrated that the extent of hybridization can be influenced by factors such as distance from an introduction source, contributing to variation in patterns among watersheds (Rubidge & Taylor 2005; Gunnell et al. 2007). In this case, pinpointing a single geographical source of introduction is impossible given that modes of white sucker introduction include unrecorded use as fishing bait and accidental hitchhiking with trout stocking efforts (Vanicek 1970; Tyus et al. 1982; Woodling 1985). Furthermore, evidence suggests that multiple introductions of the white sucker occurred over several decades (McPhee & Turner 2009) and some survey data suggest that introductions may have occurred in UCRB streams over very similar timeframes (Colorado Parks and Wildlife; data not shown). However, we cannot rule out the possibility that time since introduction combined with population size of the white sucker in a given watershed could also cause these differences in the extent of introgressive hybridization among drainages. For instance, some studies have found that the number of stocking events (Marie et al. 2012) or the number of fish
released (Hansen & Mensberg 2009) (i.e. propagule pressure) are associated with an increased level of hybridization between stocked and wild fish.

Although specific environmental conditions were not measured, they may also help explain the differences in range-wide hybridization patterns. Increased levels of hybridization have been shown to be associated with human-mediated habitat modifications, particularly those that disrupt habitat isolation among species (Riley et al. 2003; Brede et al. 2009; Kimura & Munehara 2009). For instance, the homogenizing effects of eutrophication has been implicated as the cause of increased hybridization between some fish species (e.g. Seehausen et al. 1997; Bittner et al. 2010), while logging activity and urban infrastructure development were found to influence the level of hybridization observed among coastal cutthroat and rainbow trout populations (Heath et al. 2010). Dams in particular can have major impacts on the habitats of biotic communities by changing the flow regime, temperature, and eutrophication levels and by partitioning streams into fragmented zones. Such habitat alterations have been known to impact spawning cues in several freshwater fish species. These changes can therefore dramatically affect gene flow patterns between resident fish populations, particularly if connectivity among unique habitats is required to carry out their life cycles (Schlosser 1995; Bower & Hubert 2008).

Finally, another possible driver of differences in hybridization patterns may be the presence or absence of other catostomids. The CRB is home to a number of other catostomid species including the endangered razorback, longnose, mountain, Utah, and desert suckers, all suspected to hybridize at various levels with those species included in this study (Sigler & Miller 1963; Smith 1966; Holden & Stalnaker 1975; Buth et al. 1987; Bezzerides & Bestgen 2002). Indeed, in some studies of FMS and BHS, mitochondrial genotypes of unknown origin were found that were determined to not be from WHS (e.g. Douglas & Douglas 2007a). Although we must note that none of the individuals in this study had a
mitochondrial haplotype of unknown origin, the full hybridization network may
nevertheless involve additional sucker species and be more complex than originally
thought, and further investigation of hybridization dynamics that include these other
potential players are warranted. Given all of these possibilities, it is imperative that future
research builds upon the samples already available in order to disentangle the role, if any,
of each of these possible causes.

Asymmetric mating

That a majority of the hybrids between the introduced and the two native suckers
had the mitochondrial DNA of the introduced sucker shows that hybridization bias tends to
favor male native suckers mating with white female suckers (Fig. 2.3b). Biased mating
could have a number of causes (see Wirtz 1999 for a review), but one likely cause may be
differences in relative population sizes. Unidirectional hybridization has been reported to
be common in cases where one species is much less abundant than the other such that
females of the rarer species will tend to choose males of the more abundant species (e.g.
Randler 2002; McCracken & Wilson 2011) due to what has been called the Hubbs principle
(Hubbs 1955). Chances of encounter alone will lead to a skewed ratio in observed matings
when females choose males of another species because of a lack of conspecific males. The
existence of such a scenario in this system is supported by the survey data (Fig. 2.6).
Accounting for the age of the individuals at the time of sampling (2008-2010), the years of
white sucker low abundance (1996-1999) are consistent with this explanation of why we
may have observed biased mating patterns favoring white sucker females.

Alternatively, slight differences in species spawning times combined with a longer
spawning season of males relative to females could also facilitate biased hybridization--
male native suckers may be ripe throughout the spawning season of white female suckers
while native female suckers may be ripe for a more limited period of time that does not allow for overlap with ripe white male suckers. Similarly biased hybridization has been observed in black bream (*Acanthopagrus butcheri*) and yellowfin bream (*Acanthopagrus australis*) in eastern Australia (Roberts et al. 2011). We do not currently have sufficient data on male versus female spawning periods to determine whether this is an actual driver in this system.

Additional causes could involve post-zygotic barriers such as differences in fertility, development, or mortality between crosses resulting from variation in mito-nuclear compatibility. Native female × white male hybrids may experience high rates of mortality or developmental problems while the reverse may not be true. For example, Álvarez et al. (2011) found that reciprocal crosses between salmon and trout produced offspring with drastically different mortality rates depending on which species contributed maternal genes. Such processes may also explain why hybridization bias is less pronounced in the FMS-WHS matings. It has been suggested that mtDNA may contain loci likely not involved in hybrid fitness and assortative mating such that most mitochondrial genes could function in the genetic background of a related species (Coyne & Orr 2004). If this notion applies to this system, it may very well be that while WHS mtDNA works well in the genetic backgrounds of both native species, FMS mtDNA and WHS nuDNA may be less compatible, and BHS mtDNA and WHS nuDNA even less so. Consequently, mtDNA may then also be introgressing more quickly into the native populations than nuclear DNA. Further investigation of the demographic dynamics involved through controlled cross-fertilization studies should help elucidate the exact causes of these patterns.
**Outlier loci**

Genes involved in maintaining reproductive barriers or specific adaptations should display higher levels of differentiation between populations relative to other non-adaptive regions of the genome. Because the amount of reproductive isolation between species generally increases with their phylogenetic distance, we would expect the number of highly differentiated loci to also scale with phylogenetic distance as well as the magnitude of adaptive differences between species. However, because multilocus scans rely upon measures of differentiation, our ability to detect such loci decreases when the baseline differentiation between species is already high, resulting in a reversed pattern. Consistent with this expectation, the multilocus scan detected the fewest outlier loci between the two most distant species, FMS and BHS, followed by BHS and WHS (n=3), and finally, the most loci between FMS and WHS (n=10), the most closely related species pair. A logical next step would be to identify and characterize these loci, particularly the three loci detected as outliers in both comparisons. Such analyses should provide further insight into the selection pressures that may be responsible for differentiating these species from one another.

**Implications for native species conservation**

Our results show that given enough time, hybridization with the introduced white sucker may be a substantial threat to the persistence of the native species in their pure forms and could lead to the eventual genetic swamping of the native gene pool. The flannelmouth in particular is at the highest risk of disappearing through hybridization unless selection for positive assortative mating becomes stronger. The removal programs currently in place by monitoring agencies will help curtail some of these issues. However, introgressed alleles are likely to remain in native sucker populations and persist despite
removal efforts particularly because later generation hybrids (i.e. backcrossed individuals) are difficult to correctly identify in the field.

It is also important to keep in mind that the patterns of hybridization described here represent a single snapshot of a complex and continuously changing interaction, both in time and space. The introduced white sucker has only been a part of the UCRB ecosystem for a relatively short period of time, enough for less than 10 generations in most streams. Thus, it is possible that positive assortative mating may become stronger with time, highlighting a need for continued monitoring over the next several years as well as a careful evaluation of the ultimate causes of hybridization and an assessment of which factors are contributing to the breakdown of existing reproductive barriers.

We also need to consider that the consequences of hybridization are complex and controversial, in that although hybridization can lead to the loss of genetic integrity, it can also increase the amount of genetic variation in the population, resulting in a greater potential for the population to adapt to a changing environment and selection pressures. It is therefore necessary to also examine the potential for local adaptation of the various hybrids in this system to better establish whether future conservation guidelines should focus on the protection of the two pure native species rather than on a potentially very successful hybrid swarm. Studies more directly assessing the relative fitness of known hybrid combinations across years and regions will help further characterize the trajectory of this system. Measuring hybrid fitness across genotype categories and environments can be particularly informative because it can indicate which isolating barriers prevent hybridization and the extent to which they are genetically or environmentally dependent. Future insights from these types of studies can help explain why the success of various hybrids is so variable and help predict when and if hybridization will persist. A better understanding of this variability may also help inform how and which hybrid populations
should be managed and the role hybrids may play in conservation plans—issues which have caused problems for other hybridizing species of concern (see Allendorf et al. 2001, 2004). Unfortunately, even if hybridization results in offspring with decreased or equal fitness relative to the parent species, the decline and eventual extinction of native genotypes can result, particularly if introduced species outnumber the natives (Rhymer and Simberloff 1996). For these reasons and particularly because the two native species in question are vulnerable to continued decline, studying the consequences of their hybridization has significant applied implications for their conservation.
CHAPTER 3

The effects of species traits and environmental factors on the gut and skin microbiota of wild catostomid fish

INTRODUCTION

All plant and animal species harbor many diverse microorganisms. For most vertebrates, microbial cells likely outnumber the number of host cells by at least an order of magnitude. It is therefore remarkable that we know relatively little about the role and host specificity of these microbial communities. However, many recent studies contribute to the growing recognition that microbes constitute an important component of the animal phenotype, affecting host fitness and host interactions with the environment (Fraune & Bosch 2010; McFall-Ngai et al. 2013). Host-associated microbes have been shown to be critical for a myriad of host functions including reproduction (Perlman et al. 2008; Sharon et al. 2010), development (Fraune & Bosch 2010), immune function (Lee & Mazmanian 2010), and nutrition (McCutcheon et al. 2009). For example, aphids rely on microbial symbionts in their gut for the production of critical amino acids lacking in their regular diet (Ramsey et al. 2010). Such studies indicate that microbial symbionts undoubtedly play an important role in the adaptation and evolution of the host species.

It is thus not surprising that such a tight host-microbe relationship could lead to the co-evolution of microbial communities with their hosts, such that more closely related species share more similar bacterial communities than more distantly related species (Brucker & Bordenstein 2012). A pattern of microbial communities mirroring host phylogeny has been observed in a few eukaryotic systems including pines (Redford et al. 2010), beetles (Andert et al. 2010), termites (Noda et al. 2007), mammals (Ley et al. 2008), and hominids (Ochman et al. 2010). One recent study of Nasonia wasps has even shown that gut microbes may play a role in speciation processes by causing hybrid mortality.
(Brucker & Bordenstein 2013). More direct tests such as reciprocal transplants between taxa (Rawls et al. 2006) and gene knock-out experiments (Ley et al. 2005) further support the idea that the host (or rather host genotype) selects for certain types of bacteria. The proliferation of such studies suggests that host genotype is important in determining their microbial communities. However, gut microbiota have also been shown to consistently co-vary with age in comparisons among humans (Yatsunenko et al. 2012), and diet in comparisons among humans and animal species (Muegge et al. 2011; Walter & Ley 2011; Delsuc et al. 2013).

A great deal of microbiome research has focused on the gut microbial community, which has recently led to a lot of exciting results and generalizable patterns, but the importance of other environments has not gone unnoticed. A growing body of research indicates that factors other than host traits are important in determining the community composition of some of these of other body sites. For instance, whereas the gut community of individuals may be significantly shaped by host traits such as diet (see Ley et al. 2008; Chandler et al. 2011), evidence suggests that external sites such as skin may be more prone to colonization by microbes present in the surrounding environment. In corals, for instance, changing environmental conditions can cause rapid shifts in their epithelial bacterial community (Koren & Rosenberg 2006; Rosenberg et al. 2007), and within humans, skin microbial composition is significantly affected by environmental factors such that individuals exposed to more similar environmental sources also tend to share more similar skin communities (Song et al. 2013). However, other research on hydra (Fraune & Bosch 2007) and amphibians (McKenzie et al. 2012; Kueneman et al. 2013) demonstrate that skin microbial communities still exhibit species specificity even when hosts are exposed to identical or similar environmental conditions. It is becoming more apparent that host-associated microbial communities in general are a function of host genetics, other host
traits and their interactions with the surrounding environment (see Benson et al. 2010), and that the importance of host traits (including host genotype, diet, age, and behavior) relative to environmental factors may differ across host taxa and across the different body habitats of the host.

Aquatic organisms provide a convenient system in which to further test these hypotheses surrounding the effects of host traits and the environment simultaneously. The prevailing view of fish gut colonization is that bacteria present in the surrounding environment colonize fish larvae to later become the resident communities in and on the adults (Nayak 2010), similar to how infants are seeded by microbes from maternal sources during vaginal birth (Dominguez-Bello et al. 2010). Full immersion in water starting from the egg and/or larval stage lends fish highly susceptible to colonization by aquatic bacteria. Moreover, the fish digestive tract receives water and food populated with microorganisms throughout development. Thus, it seems logical that the diversity of microorganisms present in these sources will influence the microbiota of the developing fish as some bacteria establish to become the resident gut inhabitants while others fail. Studies thus far seem to support a high level of species-specificity, often driven largely by host diet, although an effect of the environment has been detected in some studies. Bolnick et al. (pers. comm.) recently completed a study showing that intraspecific variation in the gut microbial diversity of stickleback can be explained by differences in diet (dietary specialist versus generalist). Yet a different study by Roeselers et al. (2011) found that domesticated and wild zebrafish harbor a similar gut microbiome despite coming from different source habitats, suggesting that intra-species variation of gut microbial composition may also be limited. However, the relative effects of each of these factors have rarely been evaluated using the same system, with the exception of a recent meta-analysis by Sullam et al. (2012). In their collection of data from fish spanning a broad phylogenetic scale and both marine
and freshwater environments, the authors found that the main factor shaping the gut microbial composition of fish was diet classification (i.e. herbivore, carnivore, or omnivore), followed by habitat type (i.e. marine or freshwater) (Sullam et al. 2012). However, this study did not examine different life stages, which have been shown to be important for influencing the gut communities of some mammals (Yatsunenko et al. 2012; Jami et al. 2013). How the effects of phylogeny compare with diet, or age effects is yet unknown.

Compared to gut microbes, the microbes associated with fish skin and the factors affecting them are not as well characterized. Similar to the gut community, it has been proposed that the skin colonization process depends on the bacterial composition of the surrounding water. Studies using culture-based and culture-independent methods have previously identified patterns suggestive of a large environmental effect (Horsley 1973; Hansen & Olafsen 1999; Wilson et al. 2007; Smith et al. 2009). However, it also seems likely that because the epidermis of fish acts as a protective barrier with known immunological properties (Hjelmeland et al. 1983; Shephard 1994; Subramanian et al. 2008), a unique relationship could exist between skin microbiota and their fish hosts, and that these relationships have existed long enough to exhibit species-specificity. One study supports such a correlation between fish species and a unique skin microbiota (Larsen et al. 2013). However, like most other studies that have found a species-specific aspect of skin microbiota, comparisons were focused on host taxa across a wide phylogenetic range. Whether the importance of species identity is maintained on a finer phylogenetic scale is unclear.

Natural systems allow us to examine the variability of microbial community composition with natural diets and host genetic diversity, which are usually not adequately represented in laboratory settings. One group of catostomid fish in the Colorado River Basin offers an opportunity for further exploring how both genetic composition of the host
and environmental differences may affect the composition of resident microbiota in situ. The group includes three sympatric species that can be found across a wide geographic range such that differences in gut microbial communities may be compared among congeneric species and differences in skin communities may be compared among both species and environments. Benthic omnivores, these fish are both taxonomically and ecologically distinct from the more thoroughly studied salmonids and zebrafish and have been proposed as bioindicators of water quality and pollutants in the environment (Munkittrick & Dixon 1989). Thus knowledge of their various microflora may also help us design future studies towards understanding the effects of pollutant acquisition from food sources or environmental exposure to pollutants on host health.

**Study system**

The Colorado River Basin (CRB) spans across six western states and is home to a diversity of freshwater fish, including several in the family Catostomidae (colloquially, suckers). Three species in particular are widespread: the bluehead sucker, *Catostomus discobolus*, flannelmouth sucker, *C. latipinnis*, and the white sucker, *C. commersoni*. Phylogenetic analysis shows that although these species are taxonomically in the same genus, they are fairly divergent genetically (8-12% mtDNA divergence) (Douglas & Douglas 2007a; McDonald *et al.* 2008; Doosey *et al.* 2010) with the bluehead sucker being the most divergent species of the group (Fig. 3.1). Because all three species are omnivorous benthic feeders, they overlap significantly in their diets. However, their mouth morphologies significantly differ (Fig. 3.1), and gut content studies indicate that the bluehead sucker in particular exhibits specific feeding habits (Vanicek 1967). Thus, we expect the bluehead sucker to harbor a gut community distinct from the other two catostomids for two reasons: 1) the bluehead is the most distantly related species in this group and 2) the bluehead
sucker has the most distinct diet, which has been shown to be a strong driver of gut community composition. The three species also share similar mating behavior, timing, and preferences for spawning habitat. Offspring therefore likely develop in similar habitats, allowing us to test whether environmental source is a strong driver of skin microbial community composition.

The aim of this study was to investigate the natural diversity and composition of gut and skin microbes across these three sympatric closely-related fish species in order to address two main questions. First, are either or both the gut and skin microbial communities species-specific, different compositionally depending on the environments they inhabit, or rather influenced differently by host traits and the environment? Second, how does the gut microbial community vary from the juvenile to adult age classes? Ontogenetic shifts in diets between these life stages are widespread among fish species (Werner & Gilliam 1984). In humans, the introduction of solid foods between infancy and childhood has been proposed to be a cause in the observed changes in the gut microbiota (Koenig et al. 2011; Yatsunenko et al. 2012). Thus, we might expect fish gut microbial communities to also change substantially during this process.
Figure 3.1. Phylogeny and mouth morphologies of the three sucker species. Phylogeny is based on mitochondrial sequence divergence as described in Doosey et al. (2009). Mouth morphologies are one of the features used to identify these three species in the field. The bluehead sucker has a prominent hard ridge, which the other two species lack. The flannelmouth has a relatively fleshy mouth with enlarged, lower lips compared to the bluehead and white suckers.

MATERIALS AND METHODS

Sampling and data acquisition

In the Fall of 2012, fish were sampled from three locations in the Colorado River Basin, near Grand Junction, CO. These locations were chosen because they share similar stream characteristics and contain the same species of fish, yet are distant enough from each other that individuals likely constitute separate populations, representing two subdrainages. Two of the sites are fish passageways (Redlands fish ladder, located on the Gunnison River and Government Highline fish ladder, located on the Colorado River). The third site is an open river location along a one-mile stretch of the Colorado River mainstem. Table 3.1 contains a summary of collection locations and numbers of individuals sampled along with age class information. Individuals were identified in the field to species through an inspection of morphological traits, and we assigned each fish to a life stage of juvenile or adult based on approximate length. Because these three species naturally hybridize, and
some hybrids are hard to distinguish from pure species morphologically, fin clips were taken so that species could be identified through genetic analysis.

In the fish ladders, individuals were trapped over the course of a 24-hour period in the passageways where they were held until sampling. In the stream location, fish were captured through electro-shocking, netted, and held in holding tanks containing river water prior to sampling. After individuals were rinsed thoroughly with sterile water, skin samples were collected by brushing a sterile cotton-tipped swab (BBL CultureSwab) along a ventral portion of the body. Fecal samples were collected through manual expression of fecal material into sterile microcentrifuge tubes. Although this method precludes the need for sacrificing the animal in order to obtain a gut sample, it can also lead to variable levels of success in expressing samples from individuals. This was particularly true for those fish that were held in the fish passageways, possibly because they were held for a length of time that allowed for fecal release before being processed. We were only able to collect fecal samples from one fish at the fish ladders, and therefore not able to conduct analyses testing the effects of collection location on fecal composition. A total of 68 skin samples and 39 fecal samples were collected, representing three species and two age classes. 50mL of water was also collected from each of the three sites for comparison. All samples were transferred to a -80°C freezer for storage until DNA extraction. Samples were collected under permits obtained from the Colorado Parks and Wildlife and the University of Colorado IACUC.
Table 3.1. Sample types and numbers of each species sampled at each location. Sample sizes of adults (A) and juveniles (J) of each species are given in parentheses. BHS=bluehead sucker, FMS=flannelmouth sucker, WHS=white sucker.

<table>
<thead>
<tr>
<th>SampleType</th>
<th>Location</th>
<th>BHS</th>
<th>FMS</th>
<th>WHS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fecal</td>
<td>Colorado River</td>
<td>10 (8A,2J)</td>
<td>15 (7A,8J)</td>
<td>10 (9A,1J)</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Redlands Fish Ladder</td>
<td>0 (1A,0J)</td>
<td>1 (1A,0J)</td>
<td>0 (1A,0J)</td>
<td>1</td>
</tr>
<tr>
<td>Skin</td>
<td>Colorado River</td>
<td>9</td>
<td>13</td>
<td>4</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Government Highline Ladder</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Redlands Fish Ladder</td>
<td>5</td>
<td>16</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>Water</td>
<td>Colorado River</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Government Highline Ladder</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Redlands Fish Ladder</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>29</td>
<td>46</td>
<td>15</td>
<td>93</td>
<td></td>
</tr>
</tbody>
</table>

**DNA extraction and sequencing**

Prior to DNA extraction, water samples were filtered using a 0.2 micron Nalgene Analytical Test Filter (Thermo Scientific), and swabs of the fecal samples were taken. DNA extraction and PCR amplification of the skin, fecal, and water samples followed the protocol described in Caporaso et al. (2012) and can be found on the Earth Microbiome Project website (http://www.earthmicrobiome.org/emp-standard-protocols/). Briefly, samples were extracted using the PowerSoil DNA isolation kit (MoBio Laboratories, Carlsbad, CA, USA). PCR amplification was performed using the universal bacterial primers 515F and 806R, which amplifies a region of the 16S rRNA gene spanning the V3 and V5 hypervariable regions (Caporaso et al. 2011, 2012). The 806R reverse primer contained a 12 base error-correcting Golay barcode as described in Caporaso et al. (2011). PCR was performed in triplicates, which were then pooled, and quantified using Quant-IT PicoGreen. Equal concentrations of amplicons were then pooled, cleaned using the MoBio UltraClean PCR Clean-up kit, and sequenced on an Illumina HiSeq 2000 sequencer (Illumina, San Diego,
CA, USA) at the BioFrontiers Institute Next-Generation Genomics Facility at the University of Colorado, Boulder, USA.

DNA extraction of the fin clips was performed using the DNeasy Blood & Tissue kit (Qiagen, Valencia, CA). Individuals were genotyped using AFLP (amplified fragment length polymorphism) markers. Using three selective combinations of primers, AFLP markers were generated and amplified using the protocol described in McDonald et al. (2008). Fragment analysis was performed by the Nevada Genomics Center in Reno on an ABI 3130 capillary sequencer. Fragments of sizes between 60 and 400 bp were scored as present (1) or absent (0) at each locus for each individual using the program Genemapper v4.0 (Applied Biosystems Inc.)

16S sequence analysis

16S rRNA amplicon sequence processing and analyses were conducted using the Quantitative Insights Into Microbial Ecology (QIIME) v1.7.0-dev toolset (Caporaso et al. 2010). Briefly, sequences were quality-filtered using QIIME’s default parameters and remaining sequences were subsequently clustered into Operational Taxonomic Units (OTUs) using UCLUST (Edgar 2010) against the Greengenes reference database (version October 2012) using a minimum sequence identity of 97%. 96.6% of the sequences matched the reference database, and OTUs were assigned taxonomies according to Greengenes assignments. Low-abundance OTUs (<0.0005% of the total abundance) were removed according to recommendations found in Bokulich et al. (2013), resulting in 108 samples (39 fecal, 66 skin, 3 water) with sequences ranging from 2 to 283,868 per sample (mean 81255 +/- 62927 s.d.). Samples were rarified to 10,000 sequences per sample to remove biases introduced by variation in sequence abundances, also effectively removing samples with fewer than 10,000 sequences from downstream analyses. This resulted in a final dataset of
3 water samples, 54 skin samples and 36 fecal samples representing 64 individuals (Table 3.1). 16S sequence data and associated metadata for this study are available in the EMP database (http://www.microbio.me/emp).

Diversity analyses

Rarefied abundance data were used in diversity and community similarity analyses by first calculating Unifrac distances, a phylogeny-based metric (Lozupone & Knight 2005), as well as Bray-Curtis similarity, a taxon-based measure, among samples. Both unweighted and weighted Unifrac distances were used because they are able to reflect different patterns of diversity and provide different insights into community processes; the unweighted measure better represents differences due to the presence/absence of taxa whereas the weighted measure takes into account taxon abundances (Lozupone et al. 2007). Because gut and skin communities of vertebrate hosts are already known to harbor very different microbiota (e.g. Costello et al. 2009; Song et al. 2013), we expected the main separation among samples to be between sample types. After we verified this expectation through a principal coordinate analysis (PCoA) (Fig 2), we conducted all downstream analyses separately for gut and skin samples. To determine whether location, species identity, age, and interactions among these factors were significant predictors of variability in the composition of gut and skin communities, we ran a permutational multivariate analysis of variance (PERMANOVA) using PRIMER v6 (Clarke & Gorley 2006). We ran this analysis separately for each of the distance/dissimilarity matrices using the above-mentioned factors as fixed effects. To visualize patterns of compositional differences in bacterial communities among samples, we reduced the distance matrices using a principal coordinate analysis (PCoA). To further test for significant differences in community
composition among the groups of interest, we conducted an analysis of similarity (ANOSIM) using PRIMER.

Using the QIIME toolset, bacterial diversity (alpha-diversity) was measured for each sample in two ways: Faith’s Phylogenetic Diversity (PD) (Faith 1992) and Chao1. In order to avoid biases introduced by a single rarefaction process, the full dataset was rarefied ten times at a depth of 10,000 sequences per sample, alpha diversity was calculated for each iteration, and the resulting mean value was used as the final measure. To investigate how levels of alpha diversity varied among species, age groups, and among locations, we modeled the fit of a linear model with these factors as fixed effects in R (Team 2013). To visualize the relative abundances of bacterial taxa across groups, heat maps and a taxa summary were constructed. For the heatmaps, OTU abundances were condensed to bacterial classes (e.g. Actinobacteria) for each individual and visualized separately for the fecal and skin communities. For the taxa summary, OTU abundances were condensed to the genus level for each individual and then subsequently averaged across all individuals in the group of interest to create a composite representative of the group. Classes of bacteria exhibiting significantly differences in abundance among groups were identified using a Kruskal-Wallis test. P-values were corrected for multiple comparisons using the false discovery rate (FDR) method.

Identifying probable sources of microbial taxa

In order to identify probable sources of fish skin microbes, we used SourceTracker (Knights et al. 2011). This program implements a Bayesian approach to estimate the proportions of a given sample likely derived from a set of pre-defined potential source communities. We first filtered the non-rarefied OTU table to exclude OTUs found in fewer than 10% of our samples as recommended in Knights et al. (2011). We used the three water
samples and the fish fecal samples as possible sources of microbes, a burnin of 1000, and rarefied sequences to 10000 per sample.

**Correlation with host genotype**

AFLP data were used to create a matrix of pairwise genetic distances among individuals using the RestDist package in PHYLIP v3.6 (Felsenstein 2005). In order to test whether the distance among the microbial communities is correlated to the genetic distance among the hosts, the AFLP-based distance matrix was compared with the Unifrac distance matrices using a non-parametric permutation-based Mantel test.

**RESULTS**

The PCoA of all samples, performed using unweighted UniFrac distances, showed clustering based on sample type, with water communities clustering closer to the skin samples than to the fecal samples (Fig. 3.2). All sample types including water samples were dominated by an abundance of Proteobacteria, which accounted for 64% of all sequences. However, water samples were much more diverse.
Figure 3.2. Clustering and alpha diversity of all samples. Samples are clustered based on principal coordinates analysis (PCoA) of unweighted Unifrac distances (a). Each point represents an individual sample and is colored by sample type. Collection locations are indicated by shapes. A boxplot shows the distribution of bacterial diversity by sample types (b). Diversity was measured as Faith’s Phylogenetic Diversity (PD). *p<0.05 after Bonferroni correction (Kruskal-Wallis test)

Fecal community

Proteobacteria was the most abundant phylum of bacteria in the fecal material of most individuals (53% of all fecal sequences; Fig. 3.3a), concordant with previous studies (Rawls et al. 2006; Nayak 2010; Sullam et al. 2012). However, those of several individuals were dominated instead by Fusobacteria, representing up to 96% of the fecal community in some individuals. Moreover, almost all Fusobacteria sequences belonged to the genus Cetobacterium (Fig. 3.4). Other phyla represented in moderate abundances included Verrucomicrobia (8%) and Firmicutes (7%), although abundances varied considerably across individuals (Fig. 3.3d).
Because we obtained 35 of the 36 fecal samples from the Colorado River sampling location, we focused on this subset to investigate the relative effects of host species identity and age group on the gut microbial community. Overall, variation among individuals was great, with unweighted Unifrac distances between any two individuals ranging from 0.43 to 0.90. Host species assignment explained a much larger proportion of this variance (PERMANOVA $R^2=0.10$, $p<0.001$) than host age ($R^2=0.02$, $p=0.626$; Table 3.2), although we detected an interaction between age and species such that the effects of age differed depending on which species was considered ($p=0.007$; Table 3.2; Fig 3.3). The explanatory power of species identity doubled and group separation was stronger when we used the weighted Unifrac or Bray-Curtis measures (Tables 3.2 and 3.3), indicating that differences in taxon abundances may be more important in distinguishing among host species than simply the differences in which bacterial taxa are present or absent. These were concordant with ANOSIM results, which also showed that fecal communities are significantly separated by host species identity ($p<0.001$), with differences between the white and bluehead attributing the most to this separation between species (ANOSIM $R=0.50$, $p<0.001$; Table 3.3). Moreover, gut community composition was much more consistent among blueheads (mean unweighted Unifrac distance=$0.52 +/- 0.04$ s.d.) than among whites ($0.63 +/- 0.08$ s.d.) or flannelmouths ($0.67 +/- 0.11$ s.d.), which showed greater levels of inter-individual variability (Levene’s test $p<0.001$).

The large range of variability within flannelmouths can be attributed in part to the fact that there was significant subclustering by age class (Fig. 3.3c). Indeed, ANOSIM results indicated that within this species, the fecal communities of juveniles and adults differ significantly ($R=0.296$, $p=0.007$; Table 3.3). As with PERMANOVA, this result was robust to the type of distance measure chosen. Low sample sizes of white and bluehead juveniles did not allow for similar analyses testing the effect of age for these species. We
should note that as we already mentioned, within-group variances were not equal among species nor among age groups (Table 3.3), and this could have also led to the finding of statistically significant differences among groups using these tests (Legendre & Legendre 1998). However, an evaluation of taxonomic differences and levels of diversity also suggests that the fecal composition of species differ significantly. Differences in the relative abundances of dominant taxa were consistent between species and between flannelmouth life-history stages even at the broad taxonomic levels of phylum and class (Fig. 3.3a). Perhaps most notably, flannelmouth adults had an abundant representation of Fusobacteria (Fig. 3.4), whereas flannelmouth juveniles had a more taxonomically enriched microbial community more similar in composition and diversity to that of blueheads (Fig. 3.3b; weighted Unifrac ANOSIM R=0.403, p=0.006 (FMS adults) vs R=0.87, p<0.001 (BHS)).

Levels of bacterial diversity followed a similar trend with significant differences among the species, with further differences between age classes within flannelmouths. The linear model of best fit included the factor species and an interaction term between species and age group (adjusted R²=0.46, p=0.0002; Table 3.4). Of the three species, blueheads had the most diverse gut community, regardless of which diversity measure we chose (PD or Chao1) (Fig. 3.3b). However, when we considered age as a grouping factor, the level of diversity of the flannelmouth juvenile gut was comparable to that of blueheads, whereas flannelmouth adults harbored a much more depauperate gut community (Fig. 3.3b).
a) 

<table>
<thead>
<tr>
<th></th>
<th>BHS</th>
<th>FMS</th>
<th>WHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>J</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Acidimicrobia (Actinobacteria)
Actinobacteria (Actinobacteria)*
Bacteroidia (Bacteroidetes)
Flavobacteria (Bacteroidetes)
Oscillatoriophycideae (Cyanobacteria)*
Bacilli (Firmicutes)
CK-1C4-19 (Firmicutes)
Clostridia (Firmicutes)
Fusobacteria (Fusobacteria)*
Alphaproteobacteria (Proteobacteria)*
Betaproteobacteria (Proteobacteria)*
Gammaproteobacteria (Proteobacteria)*
Mollicutes (Tenericutes)
Verrucomicrobiae (Verrucomicrobia)*

b) 

Box plot showing bacterial diversity (PD) for different species (BHS, FNS, WHS) and age classes (Adult, Juvenile).
Figure 3.3. Main factors affecting the skin bacterial community composition. A heatmap shows the relative abundances of bacterial classes (phylum in parentheses) present in fecal communities across species (a). For each sample type, OTUs with an average relative abundance of greater than 0.1% of total sequences were summarized to the bacterial class level. Taxa found to be significantly different in abundance among blueheads (BHS), flannelmouth (FMS) juveniles (J), flannelmouth adults (A), and white suckers (WHS) are bolded and starred (Kruskal-Wallis test, FDR corrected p<0.05). A boxplot shows the distribution of bacterial diversity in fecal samples by species and age class (b). Diversity was measured as Faith's Phylogenetic Diversity (PD). Skin communities are clustered based on a PCoA of unweighted (c, main panel) and weighted (c, right panel) UniFrac distance matrices. Each point represents an individual skin sample. Samples are colored by the species identity and symbols indicate the age class of the fish host. The percentage of the variation explained by the plotted principal coordinates is indicated on the axes.
Table 3.2. Results of PERMANOVA describing how variation in distances is attributed to different factors. Factors of significant effect (<0.05) are bolded.

<table>
<thead>
<tr>
<th>Sample type</th>
<th>Distance measure</th>
<th>Source of variation</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>Pseudo-F</th>
<th>p</th>
<th>%Var exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fecal</td>
<td>unweighted</td>
<td>Species</td>
<td>2</td>
<td>0.8591</td>
<td>0.4295</td>
<td>1.9921</td>
<td><strong>0.001</strong></td>
<td>9.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AgeGroup</td>
<td>1</td>
<td>0.194</td>
<td>0.194</td>
<td>0.8997</td>
<td>0.626</td>
<td>2.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Species x AgeGroup</td>
<td>2</td>
<td>0.7054</td>
<td>0.3527</td>
<td>1.6358</td>
<td><strong>0.007</strong></td>
<td>8.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>30</td>
<td>6.4686</td>
<td>0.2156</td>
<td></td>
<td></td>
<td>75.22</td>
</tr>
<tr>
<td></td>
<td>weighted</td>
<td>Species</td>
<td>2</td>
<td>1.0141</td>
<td>0.5071</td>
<td>6.1464</td>
<td><strong>0.001</strong></td>
<td>18.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AgeGroup</td>
<td>1</td>
<td>0.2157</td>
<td>0.2157</td>
<td>2.6142</td>
<td>0.053</td>
<td>3.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Species x AgeGroup</td>
<td>2</td>
<td>0.7106</td>
<td>0.3553</td>
<td>4.3068</td>
<td><strong>0.005</strong></td>
<td>12.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>30</td>
<td>2.4748</td>
<td>0.0825</td>
<td></td>
<td></td>
<td>44.33</td>
</tr>
<tr>
<td></td>
<td>Bray-Curtis</td>
<td>Species</td>
<td>2</td>
<td>17000</td>
<td>8500</td>
<td>5.736</td>
<td><strong>0.001</strong></td>
<td>18.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AgeGroup</td>
<td>1</td>
<td>2284</td>
<td>2284</td>
<td>1.5412</td>
<td>0.18</td>
<td>2.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Species x AgeGroup</td>
<td>2</td>
<td>8715</td>
<td>4358</td>
<td>2.9405</td>
<td><strong>0.003</strong></td>
<td>9.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>30</td>
<td>44457</td>
<td>1482</td>
<td></td>
<td></td>
<td>49.60</td>
</tr>
<tr>
<td>Skin</td>
<td>unweighted</td>
<td>Location</td>
<td>2</td>
<td>1.4049</td>
<td>0.7025</td>
<td>2.8755</td>
<td><strong>0.001</strong></td>
<td>9.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Species</td>
<td>2</td>
<td>0.5672</td>
<td>0.2836</td>
<td>1.1609</td>
<td>0.071</td>
<td>3.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Location x Species</td>
<td>3</td>
<td>0.7934</td>
<td>0.2645</td>
<td>1.0826</td>
<td>0.148</td>
<td>5.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>46</td>
<td>11.237</td>
<td>0.2442</td>
<td></td>
<td></td>
<td>76.26</td>
</tr>
<tr>
<td></td>
<td>weighted</td>
<td>Location</td>
<td>2</td>
<td>0.6290</td>
<td>0.3145</td>
<td>3.887</td>
<td><strong>0.003</strong></td>
<td>11.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Species</td>
<td>2</td>
<td>0.2281</td>
<td>0.1140</td>
<td>1.4093</td>
<td>0.205</td>
<td>4.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Location x Species</td>
<td>3</td>
<td>0.2699</td>
<td>0.09</td>
<td>1.1117</td>
<td>0.325</td>
<td>4.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>46</td>
<td>3.7221</td>
<td>0.0809</td>
<td></td>
<td></td>
<td>65.88</td>
</tr>
<tr>
<td></td>
<td>Bray-Curtis</td>
<td>Location</td>
<td>2</td>
<td>25412</td>
<td>12706</td>
<td>6.2086</td>
<td><strong>0.001</strong></td>
<td>19.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Species</td>
<td>2</td>
<td>5514</td>
<td>2757</td>
<td>1.3472</td>
<td>0.15</td>
<td>4.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Location x Species</td>
<td>3</td>
<td>6836</td>
<td>2279</td>
<td>1.1135</td>
<td>0.302</td>
<td>5.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residual</td>
<td>46</td>
<td>94139</td>
<td>2047</td>
<td></td>
<td></td>
<td>71.37</td>
</tr>
</tbody>
</table>
Table 3.3. Results of an analysis of similarity (ANOSIM) for significant grouping factors of fecal and skin microbial communities. Only those comparisons with sample sizes high enough for statistical testing are included. Groupings that show significant separation (R>0.25 and p<0.05) are bolded. For those comparisons showing significant grouping (p<0.05), the results of a Levene test of equality of variances are shown. Values from the same analysis using weighted (wtd) Unifrac measures are shown in parentheses. BHS=bluehead sucker, FMS=flannelmouth sucker, WHS=white sucker. COL=Colorado River open site, RFL=Redlands Fish Ladder, GH=Government Highline Fish Ladder.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Group</th>
<th>ANOSIM R value (wtd)</th>
<th>p (wtd)</th>
<th>Levene statistic</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fecal</td>
<td>All species</td>
<td>0.155 (0.255)</td>
<td>0.001 (0.002)</td>
<td>30.4</td>
<td>2.57E-12</td>
</tr>
<tr>
<td></td>
<td>BHS,FMS</td>
<td>0.085 (0.221)</td>
<td>0.114 (0.014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BHS,WHS</td>
<td>0.511 (0.626)</td>
<td>0.001 (0.001)</td>
<td>13.0</td>
<td>5.04E-04</td>
</tr>
<tr>
<td></td>
<td>FMS,WHS</td>
<td>0.04</td>
<td>0.226</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age Group</td>
<td>-0.093</td>
<td>0.912</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fecal (FMS)</td>
<td>Age Group</td>
<td>0.296 (0.87)</td>
<td>0.007 (0.001)</td>
<td>15.0</td>
<td>0.000296</td>
</tr>
<tr>
<td>Skin</td>
<td>All Locations</td>
<td>0.668 (0.461)</td>
<td>0.001 (0.001)</td>
<td>2.6</td>
<td>0.0759</td>
</tr>
<tr>
<td></td>
<td>COL, RFL</td>
<td>0.692 (0.419)</td>
<td>0.001 (0.001)</td>
<td>0.7</td>
<td>0.394</td>
</tr>
<tr>
<td></td>
<td>COL, GH</td>
<td>0.403 (0.686)</td>
<td>0.003 (0.001)</td>
<td>4.4</td>
<td>0.0367</td>
</tr>
<tr>
<td></td>
<td>RFL, GH</td>
<td>0.845 (0.402)</td>
<td>0.001 (0.002)</td>
<td>4.4</td>
<td>0.0367</td>
</tr>
<tr>
<td></td>
<td>Species</td>
<td>-0.011</td>
<td>0.597</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BHS, FMS</td>
<td>-0.027</td>
<td>0.624</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BHS, WHS</td>
<td>0.036</td>
<td>0.293</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FMS, WHS</td>
<td>-0.009</td>
<td>0.522</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skin (RFL)</td>
<td>Species</td>
<td>0</td>
<td>0.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BHS, FMS</td>
<td>0.011</td>
<td>0.403</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skin (COL)</td>
<td>Species</td>
<td>0.077</td>
<td>0.148</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BHS, FMS</td>
<td>0.158</td>
<td>0.045</td>
<td>1.8</td>
<td>0.185</td>
</tr>
<tr>
<td></td>
<td>BHS, WHS</td>
<td>0.23</td>
<td>0.088</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FMS, WHS</td>
<td>-0.238</td>
<td>0.981</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.4. Summary of the bacterial composition of each sample type. Genera abundant at a relative abundance of >0.1% of the total dataset are shown. Fecal samples were split into groups according to species and age groups; skin and water samples were divided into groups according to sampling locations. Each bar represents the average proportions across all individuals in the group. The number of individuals in each group is shown above each bar. Classifications indicate the genus of the bacteria followed by their phylum classification in parentheses; for unclassified taxa, family classification is also provided.

Table 3.4. Summary of the best fit linear model for predicting the bacterial diversity of fecal and skin communities. The response of phylogenetic diversity (PD) was modeled using host species identity (Species), age status of the host (AgeGroup), and the location of sampling (Collection Site).

<table>
<thead>
<tr>
<th>Response</th>
<th>Best model terms</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Pr(&gt;F)</th>
<th>Full model adj R² (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>fecal PD</td>
<td>Species</td>
<td>1474.21</td>
<td>737.1</td>
<td>10.6062</td>
<td>0.00033</td>
<td>0.4567 (0.0002)</td>
</tr>
<tr>
<td></td>
<td>Species x AgeGroup</td>
<td>917.56</td>
<td>305.85</td>
<td>4.4009</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Residuals</td>
<td>2084.92</td>
<td>69.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>skin PD</td>
<td>Collection Site</td>
<td>6.2673</td>
<td>3.13365</td>
<td>31.024</td>
<td>1.97E-09</td>
<td>0.5598 (4.22E-09)</td>
</tr>
<tr>
<td></td>
<td>Species</td>
<td>0.9455</td>
<td>0.47274</td>
<td>4.6802</td>
<td>0.0138</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Residuals</td>
<td>4.9494</td>
<td>0.10101</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Skin community**

Overall, Proteobacteria, mostly of the class Gammaproteobacteria, was the predominant phylum colonizing the external surface of fishes (Fig. 3.4, Fig. 3.5a), representing on average 57% of the skin community. However, similar to the fecal communities, there was substantial variation among individuals in the composition of skin communities, with unweighted Unifrac distances between individuals ranging from 0.52 to 0.82. In contrast to the fecal dataset, sampling location rather than the identity of the host species explained the largest portion of this variation (PERMANOVA $R^2=0.095$, $p<0.001$; Table 3.2). ANOSIM results indicated that there is a high level of separation among groups ($R>0.40$ in all pairwise comparisons, $p<0.005$; Table 3.3). The bacterial diversity of skin communities was also best predicted by locality (Table 3.4; Fig. 3.5c), with those from the Redlands location having the greatest diversity (Fig. 3.5b). However, we also detected a small effect of species (Table 3.4), with blueheads showing the greatest level of diversity.

Because the largest amount of variation was explained by sampling location, we analyzed locations separately to determine whether we could detect evidence for species specificity or differences between age groups within sites. Neither analysis indicated that there was significant clustering by species or age groups (Tables 3.2 and 3.3). Also in contrast to the fecal dataset, group separation was weaker when we used the weighted Unifrac metric, indicating that differences among locations are being driven more by differences in the presence/absence of particular species rather than differences in taxon abundances. Thus, we identified the taxa differentially represented in the three locations using a Kruskal-Wallis test, and investigated the potential sources of taxa driving the differences by examining the similarities between skin and water samples. We found that the skin communities of fish sampled from a given location did not share a significantly
greater number of bacterial taxa nor higher overall similarity in composition with the water sample from the same location (Fig. 3.6a and b).

Using SourceTracker, we found that the skin of fish from the Colorado River open location had a much greater abundance of microbes derived from fish feces than those from either of the fish ladders (Fig. 3.6c and d; Student’s t-test p<0.001). These skin communities were also dominated by a notably larger abundance of Acinetobacter than those from the fish ladders (Fig. 3.4). Source water did not seem to be a significant contributor to the skin communities of the fish sampled from the same location, although RFL water was found to be a highly probable source of a significant proportion (45%) of the skin samples from that location (Fig. 3.6c).
a) 

<table>
<thead>
<tr>
<th></th>
<th>COL</th>
<th>GH</th>
<th>RFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphs</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

unclassified (Actinobacteria)
Actinobacteria (Actinobacteria)*
Bacteroidia (Bacteroidetes)
Flavobacteria (Bacteroidetes)*
Sphingobacteria (Bacteroidetes)
Bacilli (Firmicutes)*
Fusobacteria (Fusobacteria)*
Alphaproteobacteria (Proteobacteria)
Betaproteobacteria (Proteobacteria)
Deltaproteobacteria (Proteobacteria)*
Gammaproteobacteria (Proteobacteria)*
Deinococci (Thermi)
Verrucomicrobiae (Verrucomicrobia)*
Pedosphaerae (Verrucomicrobia)

b) 

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>20</th>
<th>40</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Bacterial Diversity (PD)

ColoradoRiver
GoHigline
Redlands

Sampling Location

*
Figure 3.5. Main factors affecting the skin bacterial community composition. A heatmap shows the relative abundances of bacterial classes (phylum in parentheses) present in skin communities across sampling locations (a). For each sample type, OTUs with an average relative abundance of greater than 0.1% of total sequences were summarized to the bacterial class level. Taxa found to be significantly different in abundance among the Colorado River (COL), Government Highline (GH), and Redlands (RFL) sampling locations are bolded and starred (Kruskal-Wallis test, FDR corrected p<0.05). A boxplot shows the distribution of bacterial diversity in skin samples by sample location (b). Diversity was measured as Faith’s Phylogenetic Diversity (PD). *p<0.05 after Bonferroni correction (Kruskal-Wallis test). Skin communities are clustered based on a PCoA of unweighted (c, main panel) and weighted (c, right panel) UniFrac distance matrices. Each point represents an individual skin sample. Samples are colored by sampling location and symbols indicate the species identity of the fish host. The percentage of the variation explained by the plotted principal coordinates is indicated on the axes.
Figure 3.6. Comparison of similarities between skin and water samples from the same versus different locations (a-b) and source environment proportions for all skin samples estimated using the 36 fecal samples and each water sample as source environments (c-d). Similarity was assessed as the number of shared OTUs (a) and unweighted Unifrac distances (b). Bar charts show estimates of the mean proportions derived from likely sources for each sample using 100 draws from Gibbs sampling (d) and averaged values for samples from the same location (c). COL=Colorado River open site, GH=Government Highline Fish Ladder, RFL=Redlands Fish Ladder. *p<0.001 after Bonferroni correction (Kruskal-Wallis test).

Correlation with host genotype

Because the species identity of the host was found to be a significant predictor of the composition of the fecal microbial community, we used a Mantel test to determine whether the genotype of the host was also a significant factor by comparing the unweighted Unifrac distances among gut communities to a distance matrix based on the host’s AFLP profile. We
filtered this analysis to include only adult individuals as age group was also found to be an important factor. The results of the Mantel test showed that there is a significant correlation between the two distance matrices (Table 3.5). Similarly, we also tested whether distance among host genotypes was significantly correlated with the distance among skin communities, and found that in contrast to the fecal dataset, no significant relationship exists between the two. Moreover, we found that distance among skin communities was not significantly correlated with the distance among gut communities.

Table 3.5. Results from Mantel tests of correlation between distance matrices (DM) generated from host genotype (aflp) and microbial sequence data (fecal or skin). 1000 permutations were performed for each analysis to calculate the significance value (p). The numbers of individuals for which data of both types were available are shown (n). COL=Colorado River open site, RFL=Redlands Fish Ladder site.

<table>
<thead>
<tr>
<th>DM1</th>
<th>DM2</th>
<th>n</th>
<th>Mantel r statistic</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>fecal (adults)</td>
<td>aflp</td>
<td>22</td>
<td>0.22718</td>
<td>0.001</td>
</tr>
<tr>
<td>skin (COL)</td>
<td>aflp</td>
<td>25</td>
<td>0.03247</td>
<td>0.696</td>
</tr>
<tr>
<td>skin (RFL)</td>
<td>aflp</td>
<td>21</td>
<td>-0.15444</td>
<td>0.253</td>
</tr>
<tr>
<td>fecal</td>
<td>skin</td>
<td>28</td>
<td>0.16458</td>
<td>0.123</td>
</tr>
</tbody>
</table>

DISCUSSION

The undeniable role of microbial communities in host development, nutrition, immunity, and general health has led to a growing interest in understanding the genetic, environmental, ecological, and evolutionary determinants of animal-associated microbial composition and the functions these communities can perform. Yet until relatively recently, most research on microbial communities relied on culture-based methods, which we know vastly underestimates microbial diversity, and a majority of published studies, including culture-independent studies, have focused on the gut microbiota. These studies have
collectively been able to shed light on the types of factors that shape the composition and structure of host-associated microbial communities, showing that these communities can vary across species (e.g. Ochman et al. 2010), diets (e.g. Ley et al. 2008; Faith et al. 2011), developmental stages (e.g. Yatsunenko et al. 2012; Jami et al. 2013), and environmental conditions (e.g. Mouchet et al. 2012; Sullam et al. 2012). However, a better understanding of when and in which types of situations each of these factors are important and an assessment of the relative importance of each factor is necessary.

Using high-throughput sequencing technologies, we were able to further evaluate determinants of both the fecal and skin microbial communities of fish species by studying three congeneric fish in their natural habitat. In line with previous studies (Cahill 1990; Ringø et al. 1995; Smith et al. 2007; Larsen et al. 2013), we found that both the gut and skin microbiota of fish consist of a much simpler community than the surrounding water (Fig. 3.2b), suggesting that they are selective environments capable of harboring a unique microbial community. Overall, our findings indicate that while host diet and possibly host phylogeny play a role in shaping the gut microbial communities of fish, the environment plays a larger role than these factors in shaping their skin communities. Because the two communities were affected differently by host traits and environment, it was not surprising to find that the two micro-communities of the same individual were not significantly correlated. From a methodological perspective, our study also demonstrates the utility of non-destructive sampling methods. All previous studies of the fish gut microbiome have opted to fully remove the intestine or GI tract of fish; however, results reported here are similar to those reported for other fish species and indicate that fecal material is representative of the fish gut in terms of bacterial composition. Therefore, destructive sampling or euthanization is not always necessary for the study of the fish gut microbiome,
and the use of these methods may be important especially for the study of fish species of conservation concern.

The gut community: variation among species

Our results suggest that diet is the strongest driver of the gut community in this system, with the bluehead sucker having the most distinct gut community as predicted. Blueheads have a specialized cartilaginous ridge inside the upper and lower lips that the other two species lack; blueheads use this feature to primarily scrape algae from rocks, although they are also known to feed on benthic insects and detritus and thus officially are classified as omnivores (Vanicek 1967; Joseph 1977). These relatively narrow food preferences may also explain in part why their gut microbial communities exhibit little variability between individuals, despite that they also had the most diverse gut community. Able to utilize a higher proportion of algae in their diet relative to the other two species, the bluehead’s more herbivorous diet may also explain the higher number of taxa observed since an increased bacterial diversity has recently been associated with herbivores relative to omnivores and carnivores (Ley et al. 2008).

In contrast to the bluehead, the fecal communities of flannelmouth and white suckers were less diverse, did not show a similar level of distinctness, and were quite variable between individuals. The two species showed a significant amount of overlap in their microbial communities, with several individuals of both species having an unusually high abundance of Cetobacterium (phylum Fusobacteria) (Figure 3.3). This genus of bacteria is of importance for aquaculturists because Cetobacterium have been shown to be able to produce vitamin B12 and have been found in the digestive tract of freshwater fish species that do not require dietary supplements of vitamin B12 such as carp (Sugita et al. 1991; Tsuchiya et al. 2008; Ni et al. 2012). Whether this is a role provided for flannelmouth
and white sucker individuals by Cetobacterium is unknown, but their high abundance may be suggestive of such a role and further investigation is warranted.

We cannot rule out the possibility that these differences in gut microbial composition may not be driven by diet, but instead by genetic differences among species, since the bluehead sucker is the most distantly related species in the group with the flannelmouth and white suckers more closely related to each other (although not sister taxa). However, that the fecal communities of juvenile flannelmouths were significantly different from that of adults and more similar to the bluehead community (discussed in more depth in the next section) indicates that phylogenetic distance alone cannot explain all the differences observed among species.

Because all samples but one were from the same location, we could not statistically assess whether a strong environmental effect was present in this study. However, supervised classification (i.e. using a training data set to infer the value for a given sample) of the one individual sampled at the Redlands location using a random forests classifier (Breiman 2001) correctly classified this individual as its true species based on its fecal microbial composition. This suggests that environmental microbes do not have an effect larger than that imposed by species and/or diet.

If these patterns are indeed a reflection of diet, the high level of similarity between the white sucker and native flannelmouth adults is of further interest in a conservation context. The white sucker is an introduced fish to this river basin, and has already been implicated as a contributing factor to native species declines, including the bluehead and flannelmouth suckers. That they share a similar gut community may be an indication that a greater potential for food competition exists between these two species. Previous studies show that the flannelmouth and white suckers exhibit greater overlap in microhabitat preferences than either do with the bluehead sucker, which may explain these results as
similar food types can be found in similar habitat types. These findings present another potential mechanism by which the native flannelmouth may be threatened by this introduced congener.

*The gut community: change with age*

Within flannelmouths, we found what appears to be a shift in the gut microbial community between juveniles and adults. However, there have been no reports in the research to date of a corresponding shift in food preference between the juvenile and adult stages in this species. Studies do suggest however that juvenile flannelmouth suckers tend to use shallower, lower-velocity habitats compared to adults (Bezzerides & Bestgen 2002; Bower & Hubert 2008), and juveniles and adults of other sucker species have been reported to exhibit significant differences in both feeding habitat and food types (e.g. the tahoe sucker, Marrin 1983). It is thus possible that differences in food availability or exposure to environmental microbes between habitat types rather than food preferences may be driving these differences between developmental phases.

The flannelmouth juvenile fecal composition and diversity were not only different from those of flannelmouth adults, but interestingly also more similar to those of blueheads. Childs *et al.* (1998) found a high degree of diet overlap between the bluehead and flannelmouth suckers during early life stages (larval to juvenile), with an estimated Schoener’s index of 0.91 (which ranges from 0 to 1 with 1 meaning complete overlap). Our results suggest that the bluehead feeding strategy remains relatively consistent between the juvenile and adult life stages, although low sample size did not allow for adequate statistical testing of this hypothesis. We conclude that the gut microbiota is influenced by an interaction between species and developmental stage, the nature of which may be dietary, but remains unknown.
The skin community: environmental effects

The majority of previous studies of fish skin microflora, whether of freshwater or marine species, have shown that the environment is a significant determinant of the skin community. One study of whiting fish was even able to place individuals close to their known source locations through a comparison of their skin microbiota to seawater samples (Smith et al. 2009). While we were not able to make significant ties between water samples and the fish from the corresponding location, our results suggest that environmental effects are nonetheless significant. Moreover, these effects seem to be larger than species effects on the fish skin microbiota, as samples grouped primarily by location, despite coming from different fish species.

Interestingly, skin of fish from the Colorado River open location appeared to have a much higher prevalence of microbes derived from fish feces than those from the fish ladders (Fig. 3.6), and was dominated by a comparatively high abundance of Acinetobacter (Fig. 3.4). Although this genus of bacteria has previously been reported as part of both the skin and intestinal microbiota of freshwater fish (Austin 2006), it is also commonly found in soil and sediment environments as well as being a common lab contaminant. Although we cannot determine the actual source of these bacteria, a sediment source is plausible given the behavior of these fish. Because they are benthic and tend to spend a large proportion of their time along the river bottom, sediment and other materials in the sediment such as settled fecal material may be significant environmental sources of their skin microbes. A sediment source may also explain why fish from the open river location had considerably more Acinetobacter than those from the fish ladders. Unlike river bottoms, which likely contain a fairly consistent mix of substrates, sediment, and organic material, the concrete floors of fish ladders are continuously rinsed clean by constant current. Limited contact with sediment while trapped in the ladder prior to sampling may have therefore led to their
skin communities being different from that of fish sampled directly from the stream. Unfortunately, a lack of sediment samples precludes us from determining whether the sediment from the sampled stream contains bacteria in similar abundances or composition to the sampled fish skin.

Our results are in contrast to the most recent study on fish skin microbiota conducted by Larsen et al. (2013), who found the fish skin microbial community was most significantly affected by host species identity rather than location. It is important to note however, that Larsen et al. included fish representing a much broader phylogenetic range with most being more distantly related than at the species level. We hypothesize that breadth of host phylogeny may play a role in our ability to detect the effect of species identity and that the relative importance of factors that structure microbial communities is dependent on the host taxa considered, at least for the skin community. In this current study and others covering a narrow phylogenetic range of host individuals such as that of Smith et al. (2009), genotypic or physiological differences between individuals may not have been strong enough to overcome the larger effect of the environment.

We also need to consider that because the fish species in this study share similar spawning habitats and spawning times, it is possible that eggs and larvae of all three species were reared in very similar stretches of the same stream, and may also help account for the high similarity among individuals. The composition of the rearing water during the egg and larval stages has been proposed to play a large role in determining the final assemblage of the adult skin microbiome. The importance of the initial colonization stages have similarly been recognized in humans, prompting the launch of long-term studies following the progression of the microbial community through development. Given the shorter lifespans of animal species such as fish, analogous studies pairing potential sources
of microbes with development for fish species hold the potential to be important tools in helping us better understand the colonization process.

As with other snapshot studies, we are uncertain whether the diversity of bacteria we captured in this study are reflective of truly resident microbiota or influenced by the presence of transient bacteria. Although fish were rinsed with sterile water prior to sampling, it is possible that at least a portion of the bacteria we sampled were transient. Especially for fish, continuous secretion of mucus may allow for a brief residency of many water-borne bacteria, yet prevent their true colonization; such conditions would lead to the appearance of a strong environmental effect. Future studies sampling deeper in the mucosal layers as well as sampling over multiple time points as noted earlier should help identify resident versus transient taxa.

CONCLUSION

Comprising the majority of living vertebrate species (Nelson 2006), fish are an incredibly diverse clade that represent an important place in animal evolutionary history. Covering a large array of ecological, morphological, and physiological varieties, they present a unique opportunity for exploring the evolution of host-microbe relationships. Yet in spite of these qualities, their associated microbiota are largely underexplored compared to that of mammals. However, thanks to a solid base of culture-based research combined with the recent advancement of molecular technologies that have allowed us to begin characterizing the other 99% of unculturable microbial life, we are beginning to develop a better understanding of the diversity and roles of fish microbes. We know that commensals play important roles in nutrition, defense against pathogens, and a number of other crucial functions for the host. We also know that the fish microbiome is a complex and dynamic trait, shaped by a combination of and interactions between multiple environmental and
host factors. Like many questions in evolutionary biology, the challenge is now to determine how genotype interacts with the environment to affect host fitness. The results of this study along with previous studies suggest that the relative importance of each of these factors may be context dependent. While traits of the host fish may be more important in shaping the gut community, such a large environmental effect on the skin community suggests that the composition of the seeding source is also important, and that the process of colonization of the skin may thus be more dependent on environmental determinants than the gut.

Moreover, patterns of microbial shifts with age found in human and other mammal systems appear to hold in fish as well, at least for some species. However, how these changes relate to host fitness is currently unknown. Hence, in further explorations of the fish microbiome, it will be essential to elucidate the functional changes associated with shifts in microbial structure in different environmental and host contexts and how they may work to determine host fitness, and in turn, its ecology and evolution.
CONCLUSION

Increased movement of people, animals, plants, and microorganisms around the globe has naturally led to an increase in novel interactions among life forms (Ruesink et al. 1995; Vellend et al. 2007). Numerous case studies of species introductions have shown us that the introduction of new genes into an ecosystem can have variable and unanticipated effects. However, we need to now consider not only the genes of the introduced animal but also its much larger assemblage of genomes that includes those of the microbes it harbors as well, since both have the potential to affect the fitness and health of the organism, and therefore the ecosystem (McFall-Ngai et al. 2013). Accordingly, the intersection of evolutionary genetics and microbial ecology will become more important as we start learning how to address the potential effects of such multi-tiered introductions. In this thesis, I demonstrate that both approaches can be utilized to help inform conservation projects. Here, I summarize my findings and discuss key directions for development of each of these aspects to meet future conservation goals.

Informing conservation with host genetic research

My findings suggest that if the goal is to preserve the integrity of native genotypes, the concentration of conservation efforts on the bluehead and flannelmouth suckers is warranted. I show that hybridization occurs over multiple generations between an introduced and each of the native sucker species and that the potential for gene flow among species is thus high. Introgressive hybridization may therefore be slowly destroying the genetic integrity of the two native suckers where native and introduced species are sympatric. However, this process is likely to progress more quickly for the flannelmouth
sucker, because a greater range of hybrids, representing a series of backcrosses past the F1 hybrid generation, were found between this native and the introduced white sucker.

Moreover, there appears to be varying degrees of hybridization among streams. Some streams have no or few hybrids despite supporting populations of all three species (Gelwicks et al. 2009; this thesis), while other streams such as Muddy Creek seem to support all classes of hybrids including hybrids representing genetic material from all three species (McDonald et al. 2008). Although a number of factors including white sucker introduction and occurrence history as well as stream characteristics and other environmental factors may be driving these patterns, the actual cause is currently unclear and additional studies investigating each of these factors are necessary. Moreover, fitness effects of hybridization have not yet been assessed, and yet these are crucial for informed management of this system. Thus, future directions should include a detailed assessment of hybrid fitness in a variety of environments including comparisons of interspecific reciprocal crosses. Such studies will help managers identify and develop strategies for vulnerable native populations based upon an evaluation of their genetic or environmental conditions. For example, if the genetic background of populations is found to determine hybrid survival, then management efforts should be concentrated on the removal of the non-native species. Alternatively, if environmental or habitat conditions are driving differences in hybrid survival, future efforts should focus on the appropriate management of habitat (e.g. restoration or re-establishment / maintenance of connectivity).

Informing conservation efforts with microbiome research

An analysis of the microbial communities associated with these three fish species also suggests that flannelmouths could face more detrimental consequences than blueheads through its interactions with the introduced white sucker. I found that although gut
community composition was highly variable among individuals, the native flannelmouth sucker shared a similar gut composition with the non-native white sucker. It is unclear whether this pattern was driven by genetic similarity or overlap in diet. However, if diet is indeed the cause, it means that as adults, the amount of diet overlap between the introduced white sucker and the native flannelmouth sucker is greater than between either and the bluehead sucker. This may be an indication that a greater potential for competition also exists between these two species. Given that the status of the flannelmouth may already be precarious due to extensive hybridization, this added potential source of threat should be investigated further in more detail through field studies.

Interestingly, while an ontogenetic diet shift has not previously been reported for the flannelmouth between the juvenile and adult stages, I detected a distinct change in their gut microbiota with age. Although other factors may be the cause, given that strong ties between diet and gut microbiota that have been demonstrated in other animal systems, a diet shift is a likely possible explanation for this shift in the gut community. If this hypothesis is correct, this study demonstrates the utility of microbial analyses in revealing important information about an individual’s diet and raises the question of whether we should advocate the use of microbial community analysis to complement or possibly replace gut content analyses in diet studies. Such methods may be particularly useful for rare and endangered species where individuals are not usually available for destructive sampling.

More generally, my findings indicate that while host-specific traits such as diet play a role in shaping their gut microbial communities, the environment plays a larger role than these factors in shaping their skin microbiota. With the contribution of this work to a growing collection of studies of the microbiomes of other wildlife (e.g. Ley et al. 2008; Sullam et al. 2012; Larsen et al. 2013), we are now developing a strong base to begin
exploring how microbial assemblages may be affected by the introduction of nonnative species of plants and animals.

Biologists, particularly epidemiologists, are beginning to utilize the benefits of microbiome research for other conservation areas. For example, microbial bioaugmentation is showing increasing promise as a way of protecting amphibian species from the chytrid fungus *Batrachochytrium dendrobatidis* that has devastated amphibians populations world-wide (Bletz *et al.* 2013). However, the full potential of microbiome research is yet to be realized. One logical application may be in ameliorating certain challenges in keeping endangered animals in captivity. For instance, a number of enteric diseases and conditions that affect primate species in captivity may be related to alterations of the species’ microbiome as a result of changes in diet, social conditions, or medical treatment. Successful management of species in captivity may well require managing their microbiomes.

Conversely, studies of the microbiome can, in many ways, benefit from the extensive work and foundations developed by the fields of ecology and evolutionary biology. The need to incorporate ecological and evolutionary theories and frameworks into the study of the human microbiome, for instance, is now widely recognized and has already offered valuable insights (Robinson *et al.* 2010; Gonzalez *et al.* 2011; Muegge *et al.* 2011; Costello *et al.* 2012; Fierer *et al.* 2012). However, there is still much that can be learned from the fusion of fields. For instance, natural hybrid zones have traditionally been heavily utilized for the study of speciation processes. New gene combinations and potential for evolutionary novelties resulting from hybridization have made naturally hybridizing species such as the one studied in this thesis attractive systems for the study of fitness effects and other factors that may co-vary with genetics. However, this potential has not yet been fully appreciated.
in microbiome studies. Future research should leverage these systems to further explore the relationship between host genetic composition and microbial community structure.
BIBLIOGRAPHY


R Development Core Team (2013). R: A language and environment for statistical computing.


