THE PHENOTYPIC AND GENETIC DISTRIBUTION OF THREESPINE STICKLEBACK THAT INHABIT THE WILLAMETTE BASIN, OREGON, USA

by

MARK C. CURREY

A THESIS

Presented to the Department of Biology
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Master of Science

September 2014

THESIS APPROVAL PAGE

Student: Mark C. Currey

Title: The Phenotypic and Genetic Distribution of Threespine Stickleback that Inhabit the Willamette Basin, Oregon, USA

This thesis has been accepted and approved in partial fulfillment of the requirements for the Master of Science degree in the Department of Biology by:

William Cresko	Chairperson
Matt Streisfeld	Member
Susan Bassham	Member

and

J. Andrew Berglund	Dean of the Graduate School

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

Degree awarded September 2014

THESIS ABSTRACT

Mark C. Currey

Master of Science

Department of Biology

September 2014

Title: The Phenotypic and Genetic Distribution of Threespine Stickleback that Inhabit the Willamette Basin, Oregon, USA

A key to understanding the origin and maintenance of the diversity of life is to understand how phenotypic and genetic variation is partitioned within and among populations. I characterize the spatial partitioning of phenotypic and genetic variation in an old Willamette Basin freshwater distribution of threespine stickleback (*Gasterosteus aculeatus*) and compare these results to younger populations. Phenotypic variation was measured using 14 phenotypic traits, and genetic variation was assessed using RADseq and *Stacks* software to identify tens of thousands of single nucleotide polymorphisms. The major partitioning of phenotypic and genetic variation in Oregon is along a stereotypical transition from oceanic to freshwater that has been seen in younger systems. Phenotypic and genetic variation is significantly partitioned between basin populations, and the genetic variation is geographically structured. This work suggests that parallel divergence between oceanic and freshwater forms originated before the end of the last glacial maximum.

CURRICULUM VITAE

NAME OF AUTHOR: Mark C. Currey

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
Eastern Oregon University, La Grande
Chemeketa Community College, Salem, Oregon

DEGREES AWARDED:

Master of Science, Biology, 2014, University of Oregon
Bachelor of Science, Biology, 2004, University of Oregon
Associate of Arts, 1994, Chemeketa Community College

AREAS OF SPECIAL INTEREST:

Evolutionary Genetics
Population Genetics

PROFESSIONAL EXPERIENCE:

Research Technician, University of Oregon, 2004-2011

PUBLICATIONS:

Catchen, J, S Bassham, T Wilson, **M Currey**, C O'Brien, Q Yeates, and WA Cresko. "The Population Structure and Recent Colonization History of Oregon Threespine Stickleback Determined Using Restriction-Site Associated DNA-Sequencing." *Mol Ecol* 22, no. 11 (2013): 2864–83.

Kimmel, CB, WA Cresko, PC Phillips, B Ullmann, **M Currey**, F von Hippel, BK Kristjánsson, O Gelmond, and K McGuigan. "Independent Axes of Genetic Variation and Parallel Evolutionary Divergence of Opercle Bone Shape in Threespine Stickleback." *Evolution* 66 (2012a): 419–34.

Kimmel, CB, PA Hohenlohe, B Ullmann, **M Currey**, and WA Cresko. "Developmental Dissociation in Morphological Evolution of the Stickleback Opercle." *Evol Dev* 14, no. 4 (2012b): 326–37.

Hohenlohe, PA, S Bassham, **M Currey**, and WA Cresko. “Extensive Linkage Disequilibrium and Parallel Adaptive Divergence Across Threespine Stickleback Genomes.” *Philos Trans R Soc Lond B Biol Sci* 367, no. 1587 (2012): 395–408.

McGuigan, K, N Nishimura, **M Currey**, D Hurwit, and WA Cresko. “Cryptic Genetic Variation and Body Size Evolution in Threespine Stickleback.” *Evolution* 65, no. 4 (2011): 1203–11.

McGuigan, K, N Nishimura, **M Currey**, D Hurwit, and WA Cresko. “Quantitative Genetic Variation in Static Allometry in the Threespine Stickleback.” *Integr Comp Biol* 50, no. 6 (2010): 1067–80.

Baird, NA, PD Etter, TS Atwood, **M Currey**, AL Shiver, ZA Lewis, EU Selker, WA Cresko, and EA Johnson. “Rapid Snp Discovery and Genetic Mapping Using Sequenced Rad Markers.” *PLoS One* 3, no. 10 (2008): e3376.

Kimmel, CB, W Aguirre, B Ullmann, **M Currey**, and WA Cresko. “Allometric Change Accompanies Opercular Shape Evolution in Alaskan Threespine Sticklebacks.” *Behaviour* (2008): 669–91.

Kimmel, CB, B Ullmann, C Walker, C Wilson, **M Currey**, PC Phillips, MA Bell, JH Postlethwait, and William A Cresko. “Evolution and Development of Facial Bone Morphology in Threespine Sticklebacks.” *Proc Natl Acad Sci USA* 102, no. 16 (2005): 5791–96.

Cresko, WA, A Amores, C Wilson, J Murphy, **M Currey**, P Phillips, MA Bell, CB Kimmel, and JH Postlethwait. “Parallel Genetic Basis for Repeated Evolution of Armor Loss in Alaskan Threespine Stickleback Populations.” *Proc Natl Acad Sci USA* 101, no. 16 (2004): 6050–55.

ACKNOWLEDGMENTS

This work would not have been possible without the many hours and much effort put in to processing samples and collecting phenotypic measurements: Steven James, Shira Mali, Savanah Olroyd, Erik Parker, Aimee Shultz, and Roberta Torunsky. Thank you!

I would also like to thank ODFW and NOAA with their help in obtaining stickleback collections especially Brian Bangs, Paul Olmstead, Sara Akins, Brian Cannon, Randy Wildman, and Laurie Weltcamp. Their help allowed me to cast a very large net.

The members of the excellent research community at the University of Oregon, including everyone in the Cresko Lab, your comments and support was very helpful. I also extend my gratitude to my committee for guiding me through this process. I would especially like to thank Susie Bassham and Julian Catchen for their excellent mentorship and my advisor Bill Cresko for his guidance and enthusiasm about this project and allowing me to strike a path of scientific discovery.

Finally, I thank my family. To my mother Christine for the Saturday morning phone calls in which you always kept me grounded. Theresa, Luke, and Delaney, thank you for all of your support and patience through this process and allowing me to pursue a life in scientific discovery.

This work was made possible because of generous funding support from the National Institutes of Health (R24 GM079486) and the National Science Foundation (DEB 0919090 & DEB 1038587).

For my wife Theresa and children Luke and Delaney who have endured and encouraged my pursuit of everything fishy.

TABLE OF CONTENTS

Chapter Page

Chapter Page

Chapter Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

WHY STUDY PHENOTYPIC AND GENETIC VARIATION?

Understanding how populations of organisms phenotypically and genetically diverge within and among populations – and across space and time – is key to understanding the origins and maintenance of biodiversity. Careful characterization of the distribution of phenotypic variation and inference from the patterns that were observed inspired Darwin's and Wallace's ideas of evolution and natural selection (Darwin, 1859; Wallace, 1871). Continued research in this area is leading to a better understanding of the processes that shape and mold phenotypic variation within and among populations (Bates, 1843; Darwin, 1859; Wallace, 1871; Wilson, 1971; Endler, 1990; Reznick et al., 2002; Seehausen, 2006; Steiner et al., 2007; Grant and Grant, 2008; Reimchen et al., 2013). Additionally, knowledge of how genetic variation is partitioned within and among populations is critical to the inference of evolutionary mechanisms that produce the phenotypic patterns observed in nature (Dobzhansky, 1937; Dobzhansky, 1955; Wright, 1942; Wright, 1978; Haldane, 1955; Fisher, 1958; Endler, 1986).

Research over the last century has led to a much better understanding of the evolutionary processes - such as genetic drift, migration, and natural selection - that mold and structure phenotypic and genetic variation within and among populations (Wright, 1943a; Sokal et al., 1989; Hoekstra et al., 2004; Bustamante et al., 2005; Storz, 2005). However, these types of studies have been limited to using few genetic markers - which

limits the inferences that can be made to very small subsets of the genome - and restricted to comparing recently diverged populations. It is still unclear how populations genetically diverge at the genome-wide level, and how populations phenotypically and genetically diverge over disparate time scales. The basic question that I addressed in my thesis is the following: How do populations diverge phenotypically and genomically through space and time?

To address this question I studied the threespine stickleback (*Gasterosteus aculeatus*). These fish are distributed holarctically and have oceanic, anadromous, and freshwater life history forms (Bell and Foster, 1994). Anadromous populations of stickleback have repeatedly colonized freshwater habitats, resulting in a large set of independent natural evolutionary experiments. The resulting freshwater populations often phenotypically diverge from oceanic ancestors in parallel ways. Surprisingly, at least some of the parallel phenotypic divergence observed in many of these populations appears to have a parallel genetic basis (Colosimo et al., 2005; Hohenlohe et al., 2010; Deagle et al., 2012; Jones et al., 2012).

Though there are broad parallels in how freshwater stickleback are different from marine ones, there is much phenotypic diversity among freshwater populations (Rundle, 2000; Hendry et al., 2011; Reimchen et al., 2013). These differences in freshwater forms are likely the result of localized adaptation to a diversity of physical parameters in freshwater habitat types (Lavin and McPhail, 1986; Bell and Foster, 1994; Kalbe and Kurtz, 2006; Hendry et al., 2011; Reimchen et al., 2013).

Importantly, these different forms of stickleback - anadromous, oceanic, and freshwater populations - are extant, allowing direct phenotypic and genetic comparisons

among them. Most of the work so far has focused on relatively young post-Pleistocene populations that arose since the last glacial maximum (approximately 10-15 thousand years ago), including some as recently as 50 years ago (Bell and Ortí, 1994; Cresko et al., 2004; Hohenlohe et al., 2010; Lescak et al., 2013; Olafsdottir et al., 2007; Rundle, 2000; Hendry et al., 2011; Reimchen et al., 2013; Kitano et al., 2008; Bell et al., 2004; Lucek et al., 2014; Gelmond et al., 2009). However, very little is known about freshwater populations of stickleback that are much older. Do populations of freshwater stickleback that have persisted since before the last glacial maximum also demonstrate the parallel phenotypic and genetic divergence from the oceanic form that is observed in younger populations? Additionally, how is genetic and phenotypic variation partitioned within and between these older populations and how does this compare to what has been observed in younger populations?

THE OLD AND ECOLOGICALLY DIVERSE WILLAMETTE BASIN

The Willamette River Basin on the western edge of the state of Oregon (USA) is very old and has a diversity of aquatic habitats making for an ideal system to address these questions. The basin was formed ~ 20 mya when uplift of the Coastal Mountain Range separated it from the Pacific Ocean (Wallick et al., 2007). Much of what is presently Oregon was not glaciated during the last glacial maximum, and many aquatic habitats, specifically those of the inland drainages, are much older than more northern or coastal ones (Benner and Sedell, 1997; O'Connor, 2001; Boothl et al., 2003). I hypothesized that stickleback populations that inhabit the Willamette Basin may have

been founded before the last glacial maximum and could potentially be as much as 10-15 million years old.

Present day aquatic habitats in the Willamette Basin are also very diverse (Hughes and Gammon, 1987; Waite and Carpenter, 2000) exemplified by differences in temperature and flow regime (Tague and Grant, 2004; Jefferson et al., 2007; Tague et al., 2007). Differences exist in streambed morphologies within and between the eastern and western tributaries in the basin. A generality of eastern tributaries of rivers in the Willamette Basin such as the McKenzie River is that they are rain and snowmelt fed because they flow from high elevation headwaters through deep ground water reservoirs resulting in relatively constant flow and consistent cool temperature regimes year round. These rivers flow over and through volcanic rock that creates cobbled streambeds ideal for salmonid spawning. In contrast, rivers and streams on the western side of the Willamette Basin, such as the Mary's River, originate from lower elevation coast range hills and are mainly rainfall-dependent resulting in high variability in flow and temperature regime (Pearcy, 1999). These rivers flow over and through sedimentary rock resulting in sandy and silty streambeds. Additionally, the main stem of the Willamette drainage includes side channels and discontiguous habitat (Benner and Sedell, 1997). This diversity of habitats led me to predict the presence of locally adapted stickleback populations in the basin.

THE RESEARCH PROBLEMS ADDRESSED IN THIS THESIS

The very old and very diverse Willamette Basin, therefore, provides a unique opportunity to investigate the phenotypic and genetic divergence of freshwater

populations of stickleback that predate the last glacial maximum and allows me to compare these findings to younger freshwater stickleback populations. This is the first in-depth investigation of the distribution of phenotypic and genetic variation within and between populations of threespine stickleback that inhabit the Willamette Basin.

In chapter II, I characterized the distribution of phenotypic variation between oceanic and Willamette Basin populations and within and between populations in the basin. I found that the major axis of phenotypic variation is between oceanic and basin populations and closely resembles what has been seen in younger freshwater systems that have diverged from oceanic forms. Phenotypic variation was also partitioned among populations within the basin. These results suggest that the parallel divergence of freshwater forms from oceanic forms originated before the last glacial maximum. This also suggests that freshwater stickleback that inhabit the Willamette Basin have phenotypically diverged from one another perhaps as the result of localized adaptation. Surprisingly, I also discovered the existence of phenotypically oceanic stickleback in the eastern region of the basin and a phenotypic cline in the McKenzie River.

In chapter III, I investigated the genetic structure and substructure of Oregon and the Willamette Basin stickleback populations and determined the amount of genetic divergence between these populations. I found that the major partitioning of genetic variation is between coastal and inland populations with large amounts of genomic differentiation between these populations I also found that there is substructure within the basin populations that is geographically distributed. These results suggest that basin populations are older than populations that arose since the last glacial maximum. I also genetically tested the hypothesis of a recent introduction of coastal fish that produced a

phenotypic cline in the McKenzie River. Genetic analysis is consistent with a recent exogenous introduction and indicates a genetic cline underlies the phenotypic cline in this river.

Finally, Chapter IV summarizes the results from Chapters II and III and discusses how they contribute to our understanding of how phenotypic and genetic variation is partitioned through space and time.

CHAPTER II

DISTRIBUTION OF PHENOTYPIC VARIATION IN THREESPINE STICKLEBACK (*GASTEROSTEUS ACULEATUS*) POPULATIONS FROM THE WILLAMETTE BASIN, OREGON

INTRODUCTION

The origins of biodiversity have long fascinated scientists. A central challenge for modern evolutionary study is to understand the biological mechanisms that lead to the origin and maintenance of this diversity. Key to addressing this problem is understanding how populations of organisms phenotypically and genetically diverge within and among populations - and across space and time - in response to environmental pressures (Bates, 1843; Darwin, 1859; Wallace, 1871; Wilson, 1971; Endler, 1990; Reznick et al., 2002; Seehausen, 2006; Steiner et al., 2007; Grant and Grant, 2008; Reimchen et al., 2013). Much of the historical (Bates, 1843; Darwin, 1859; Wallace, 1871) and contemporary (Seehausen, 2006; Grant and Grant, 2008; Reimchen et al., 2013) work has relied heavily on careful characterization of the distribution of phenotypic variation as the basis for inferring the existence and trajectory of phenotypic evolution and adaptation.

Parallel phenotypic change in response to similar environments is a strong indication that natural selection, rather than neutral processes, is a driver (Endler, 1977; Endler, 1986; Schluter, 2000). Biological systems exhibiting parallel evolution can be exploited to help identify adaptive traits and the genetic basis of their adaptive response (Cresko et al., 2004; Protas et al., 2006; Shapiro et al., 2006; Hohenlohe et al., 2010;

O'Quin et al., 2010; Rosenblum et al., 2010; Jones et al., 2012; Butlin et al., 2014) Those that demonstrate repeated parallel divergence are exceptionally useful for this type of work (Schluter and Rambaut, 1996; Jones et al., 2013). However, many of these studies have been limited to evolutionary scenarios of relatively recent divergence and it is still unclear how repeated parallel divergence is manifested over disparate time scales. Are there similar patterns and magnitudes of phenotypic and genetic change over large periods of time in similarly derived populations? What is the genetic basis of repeated divergence over large time scales? In many evolutionary model systems, older relic forms no longer exist making it impossible to address these questions.

Threespine stickleback (*Gasterosteus aculeatus*) are distributed holarctically and are abundantly present in oceanic, estuarine, and freshwater habitats. These small fish have oceanic, anadromous, and freshwater life history forms (Bell and Foster, 1994). Anadromous populations of stickleback have repeatedly colonized freshwater habitats, resulting in a large set of independent natural evolutionary experiments. The resulting freshwater populations often phenotypically diverge from oceanic ancestors in parallel ways. These phenotypic differences comprise (but are not limited to) loss of armor plates, changes in craniofacial morphology, and reduction or loss of pelvic structures (Bell and Foster, 1994). Surprisingly, at least some of the parallel phenotypic divergence observed in many of these populations appears to have a parallel genetic basis (Colosimo et al., 2005; Hohenlohe et al., 2010; Deagle et al., 2012; Jones et al., 2012).

Though there are broad parallels in how freshwater stickleback are different from marine ones, there is much phenotypic diversity among freshwater populations. A significant amount of work has been performed investigating different freshwater forms

including divergence between benthic and limnetic habitats (Rundle, 2000), divergence between clear and tannic water systems (Reimchen et al., 2013), and divergence between lake and stream forms (Hendry et al., 2011). These differences in freshwater forms are likely the result of localized adaptation to a diversity of physical parameters in freshwater habitat types (Lavin and McPhail, 1986; Bell and Foster, 1994; Kalbe and Kurtz, 2006; Hendry et al., 2011; Reimchen et al., 2013).

Importantly, these different forms of stickleback - anadromous, oceanic, estuarine, and freshwater populations - are extant, allowing direct phenotypic and genetic comparisons among them. Most of the work so far has focused on relatively young post Pleistocene populations that arose since the last glacial maximum (approximately 10-15 thousand years ago), in south-central Alaska (Bell and Ortí, 1994; Cresko et al., 2004; Hohenlohe et al., 2010; Lescak et al., 2013), Iceland (Olafsdottir et al., 2007), and British Columbia (Rundle, 2000; Hendry et al., 2011; Reimchen et al., 2013). Additionally, work is starting to be done with even younger populations, some of which have arisen within the last century, such as those found in Washington State (Kitano et al., 2008), mainland Alaska (Bell et al., 2004), Iceland (Lucek et al., 2014), and Middleton Island, Alaska (Gelmond et al., 2009). However, very little is known about freshwater populations of stickleback that are much older. Do populations of freshwater stickleback that have persisted since before the last glacial maximum also demonstrate the parallel phenotypic and genetic divergence from the oceanic form that we see in younger populations? And, will these freshwater populations phenotypically diverge from one another as has been observed in other freshwater stickleback systems?

The Willamette River Basin on the western edge of the State of Oregon (USA) is very old and has a diversity of aquatic habitats. Much of what is presently Oregon was not glaciated during the last glacial maximum, and many aquatic habitats, specifically those of the inland drainages, are much older than more northern or coastal ones (O'Connor, 2001; Boothl et al., 2003). We hypothesized that stickleback populations that inhabit the Willamette Basin may precede the last glacial maximum and could potentially be millions of years old. This is approximately the same order of magnitude as the divergence between the congeneric species threespine and ninespine stickleback (Aldenhoven et al., 2010) and we reasoned that these populations may provide important resources in potentially having been formed early in the history of the threespine stickleback lineage. Additionally, present day aquatic habitats in the Willamette Basin are very diverse (Hughes and Gammon, 1987; Waite and Carpenter, 2000) exemplified by differences in temperature and flow regime (Tague and Grant, 2004; Jefferson et al., 2007; Tague et al., 2007), and differences in streambed morphologies within and between the eastern and western tributaries in the basin. A generality of eastern tributaries of rivers in the Willamette Basin, such as the McKenzie River, is that they are rain and snow-melt fed because they flow from high elevation headwaters resulting in relatively constant flow and temperatures year round. In contrast, rivers and streams on the western side of the Willamette Basin, such as Mary's River, originate from lower elevation coast range hills and are mainly rainfall-dependent resulting in high variability in flow and temperature regime (Pearcy, 1999). This diversity of habitats leads us to predict the presence of locally adapted stickleback populations in the basin. This very old and very diverse basin, therefore, provides a unique opportunity to investigate the phenotypic and

genetic divergence of freshwater populations of stickleback that predate the last glacial maximum.

In this study we aimed to investigate the phenotypic distribution of a potentially old and a relatively unexplored freshwater group of stickleback populations that inhabit the Willamette Basin. We did this by analyzing phenotypic divergence between oceanic and Willamette Basin populations, divergence among regionally diverse groups of basin populations, and divergence among individual populations within the basin. Previous research on stickleback divergence suggests that much of the genetic variation important for initial adaptation to fresh water is carried by but not expressed in marine populations and predates the last glacial maximum (Schluter and Conte, 2009; Hohenlohe et al., 2012), so we expect to find broadly similar phenotypic change from the marine form in both old and very young freshwater populations. Additionally, we tested whether populations within the basin demonstrate localized phenotypic adaptation. We predicted that this local adaptation would be manifested in two different ways: First, because there are large fluvial differences between tributaries of the East and West regions of the basin, we hypothesized that there will be significant phenotypic partitioning of variation between these regions. Second, because of the large amount of variation in aquatic ecology throughout the Willamette Basin as well as very large distances between tributaries we hypothesized that there will be significant partitioning of variation between the populations within the basin.

METHODS

Geographic distribution of collections

Populations were sampled throughout the Willamette Basin, including most of the main tributaries within the basin (Figure 2.1, Table 2.1). Most of these collections were performed in 2012 and 2013. However, Riverbend, Science Factory, Reed Canyon, and Cushman Slough populations were collected prior to 2012. These populations included five from the main stem of the Willamette R. located in the central region of the basin, one from the Mollala R., three from the Santiam R., nine from the McKenzie R., three from the Middle Fork of the Willamette R. located in the East region of the basin, one population from the Luckiamute R., three from the Mary's R., and two from the Coast Fork of the Willamette R. located in the West region of the basin. Additionally, one population from Cushman Slough, located in the tidal zone of the Siuslaw R., and one population from the tidal zone of the Columbia R. were collected. These last two populations were used as comparators to the basin populations.

Collection of stickleback samples

The Willamette Basin is very large with much suitable stickleback habitat. Collections were made using minnow traps as previously described (Cresko et al., 2004; Catchen et al., 2013). Non-baited 0.635 cm mesh minnow traps were placed in aquatic locations and left for 8-24 hours. Up to 100 stickleback from each location were euthanized using MS222 and fixed in 95% ETOH. All collections followed IACUC protocol # 13-07. GPS coordinates of each location were obtained using Google Earth. Additional samples were obtained from state and federal agencies, namely Oregon

Department of Fish and Wildlife (ODFW) and National Oceanic and Atmospheric Administration (NOAA), which were conducting collections of other fish species in the Willamette and Columbia basins. These samples were euthanized and immediately fixed in 95% ETOH. These efforts resulted in a total of 29 populations obtained and used for phenotypic analysis (Figure 2.1, Table 2.1).

Sample preparation

Each fish from each collection was uniquely numbered. Both pectoral fins and the caudal fin were taken from each individual and stored at -20º C or immediately processed for DNA extraction. The soma of each individual was then fixed with 4% PFA overnight and stained with Alizarin Red for phenotypic analysis as previously described in (Cresko et al., 2004; Catchen et al., 2013).

Phenotypic data collection

Phenotypic distributions were investigated using morphological traits that have been previously inferred to be adaptive during the transition between ocean and freshwater or during adaptation to different freshwater environments. Lateral images of each fish were taken using a Nikon D70 digital camera. Ventral images were taken using an Olympus SZX16 dissecting microscope equipped with an Olympus DP71 microscope digital camera. Images were taken and processed with Olympus DP Controller version 3.3.1.292. ImageJ (Schneider et al., 2012) was used to measure all continuous traits from images taken as described above and to measure the angle of the head.

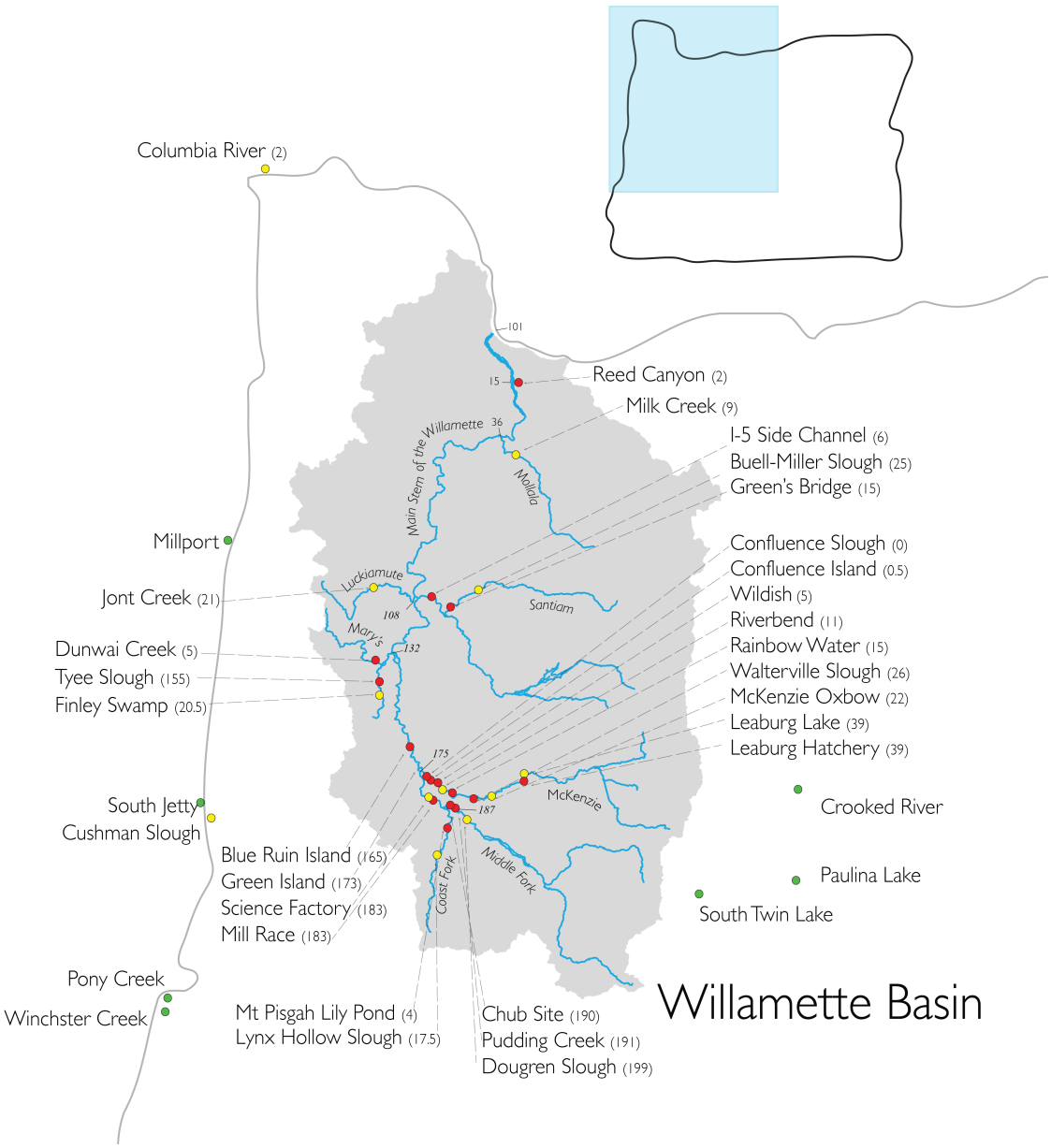

Figure 2.1. Distribution of populations used for phenotypic and genetic analysis. Yellow dots represent populations used in both analyses, red dots represent populations used only in the phenotypic analysis, and green dots represent populations only used in the genotypic analysis. The number at the confluence of the Willamette – Columbia is the number of river miles from the ocean. Numbers marking major confluences along the main stem of the Willamette indicate river miles from the Willamette – Columbia confluence. Numbers in parentheses after population names indicate river miles within each river. Populations are grouped by tributary and listed according to river mile chronologically.

Table 2.1. Collection sites with GPS coordinates, collection dates, and number of individuals used in either the genotypic or phenotypic analysis.

Population	GPS coordinates of collection site	Collection date(s)	Genotyped	Phenotyped
Blue Ruin Island	44°13’35.77"N 123° 9'15.27"W	Nov, 2012		28
Buell-Miller Slough	44°46’12.44"N 122°50'46.96"W	Aug, 2013	21	21
Chub Site	44° 1’29.00"N 122°58'29.49"W	Jul, 2012		50
Columbia River	46°14’36.04"N 123°54'9.11"W	Jul, 2012	48	60
Confluence Island	44° 7’22.07"N 123° 6'8.59"W	Jun, 2013		15
Confluence Slough	44° 7’32.21"N 123° 6'22.13"W	Jun, 2013		19
Crooked River	44°17’25.57"N 120°50'47.07"W	Aug, 2008		30
Cushman Slough	43°59’22.4″N 124°2’42.94″W	March, 2009	95	31
Dougren Slough	43°58’1.19"N 122°52'8.41"W	Sept, 2013	48	30
Dunawi Creek	44°32’55.04"N 123°18'9.66"W	Aug, 2013		30
Finley Grey Creek Swamp	44°23’44.09"N 123°20'35.42"W	Aug, 2013	48	29
Green Island	44° 8’42.02"N 123° 7'4.88"W	Oct, 2012	48	30
Green’s Bridge	44°42’32.23"N 122°58'12.74"W	Sep, 2013		30
I-5 Side Channel	44°44’11.80"N 123° 2'56.67"W	Aug, 2013		23
Jont Creek	44°46’28.32"N 123°19'5.02"W	Oct, 2012	48	30
Leaburg Fish Hatchery	44° 8’3.84"N 122°36'31.86"W	Oct, 2012		50
Leaburg Lake	44° 8’12.99"N 122°36'44.99"W	Oct, 2012	48	30
Lynx Hollow Slough	43°51’35.07"N 123° 1'24.98"W	Aug, 2013	43	30
McKenzie Oxbow	44° 3’41.29"N 122°51'10.87"W	Oct, 2012		30
Milk Creek	45°14’12.82"N 122°37'54.38"W	Oct, 2013	32	30
Mill Race	44° 2’51.59"N 123° 4'20.22"W	Previous to 2008		29
Millport Slough	44°53′14.68″N 123°59′46.20″W	Apr, 2010	68	
Mt Pisgah Lily Pond	44° 0’0.29"N 122°58'47.05"W	Aug, Dec, 2013		51
Paulina Lake	43°42’48.66"N 121°16'21.48"W	Aug, 2008	21	
Pony Creek Reservoir	43°22’12.03"N 124°15'43.52"W	Mar, 2007	70	
Pudding Creek	44° 1’15.72"N 122°57'37.85"W	Jul, 2012		39
Rainbow Water	44° 3’42.63"N 122°58'7.69"W	Aug, 2012		25
Reed Canyon	45°28’54.66"N 122°37'48.04"W	Jul, 2009		29
Riverbend	44° 4’40.97"N 123° 1'34.92"W	Various	140	41
Science Factory	44° 3’25.50"N 123° 4'30.93"W	Various	48	62
South Jetty	44° 0'7.76"N 124° 7'59.26"W	Mar, Apr, 2009	86	
South Twin Lake	43°42’38.38"N 121°46'4.53"W	Jun, 2007	50	
Tyee Slough	44°27’52.57"N 123°19'4.96"W	Aug, 2013		29
Walterville Slough	44° 4’13.64"N 122°47'54.68"W	Aug, 2012	24	50
Wildish	44° 7’0.53"N 123° 4'17.23"W	Jun, 2013		30
Winchester Creek	43°16’37.39"N 124°19'8.57"W	2007	22	
		Total	1031	982

All meristic phenotypes were collected by at least two independent researchers. Trait counts that differed in scores (e.g. differences in the number of lateral plates) between collectors were repeated until congruence was reached. Up to 30 individuals, and in some populations up to 60 individuals, were scored for all phenotypic traits analyzed.

The number of lateral plates on the left side of each fish, and number of long gill rakers of the first arch on the right side, were each quantified. Right sides were used for gill raker counts to preserve the left side of each fish for further phenotypic analysis.

Two different measures were conducted to capture aspects of body size. Standard length was measured from the tip of the pre-maxillary to the posterior end of the caudal peduncle. This measure was used to standardize all other linear measures. Body depth was measured from the mid-point of the pelvic joint articulating element to a point along the dorsal edge, anterior to the second dorsal spine, determined by drawing a line perpendicular to the long body axis of the fish.

Dorsal and pelvic spine length and different aspects of the pelvic structure were measured. The length of the second dorsal and the left pelvic spine were measured from the inflection point of the socket to the tip of the spine. The ascending process of the pelvic structure was measured from the mid-point of the pelvic joint articulating element to the most dorsal tip of the process. In cases where the ascending process ends in multiple cusps, the posterior most-cusp was used. The length of the pelvic structure was measured from the most anterior point of the left anterior process to the tip of the left posterior process. Pelvic structure width was measured as the distance between the two pelvic spine joints measured from the medial side of each joint.

Six measurements (five linear and one angular) were used to capture head morphology differences between populations. Head angle was measured by drawing a line starting at the midline of the fish directly under the last dorsal fin ray to the tip of the nasal bone and another from this point to the dorsal-most point of the skull directly above the center of the eye. The resulting angle was measured. Eye socket diameter was measured from slightly posterior of the dorsal-most aspect of the socket to the slightly anterior to the ventral-most aspect of the socket, approximately from 5° to 185°. Jaw length was measured from the anterior-most tip of the pre-maxilla to the joint between the pre-maxilla and the maxilla. The dorsal cranial length was measured from the anterior most tip of the nasal bone to the posterior most tip of the frontal. The opercle bone roughly resembles a triangle. The top of the triangle, and top of the opercle, is composed of a joint (J) where the bone hinges and pivots. The other two points of the triangle are a ventral point (V) and a posterior point (P). Linear measurements were taken from the joint to the ventral point (JV) and from the joint to the posterior point (JP; Figure 2.2).

Size standardization and data filtering

Linear measurements scale with the overall size of the organism. This scaling between size and the trait of interest needs to be accounted for to investigate the differences in traits free from scaling. To do this we employed a method described by Reist (Reist, 1985) and used in other morphological studies of threespine stickleback, (Reimchen and Nosil, 2006; Reimchen et al., 2013) that was calculated as follows.

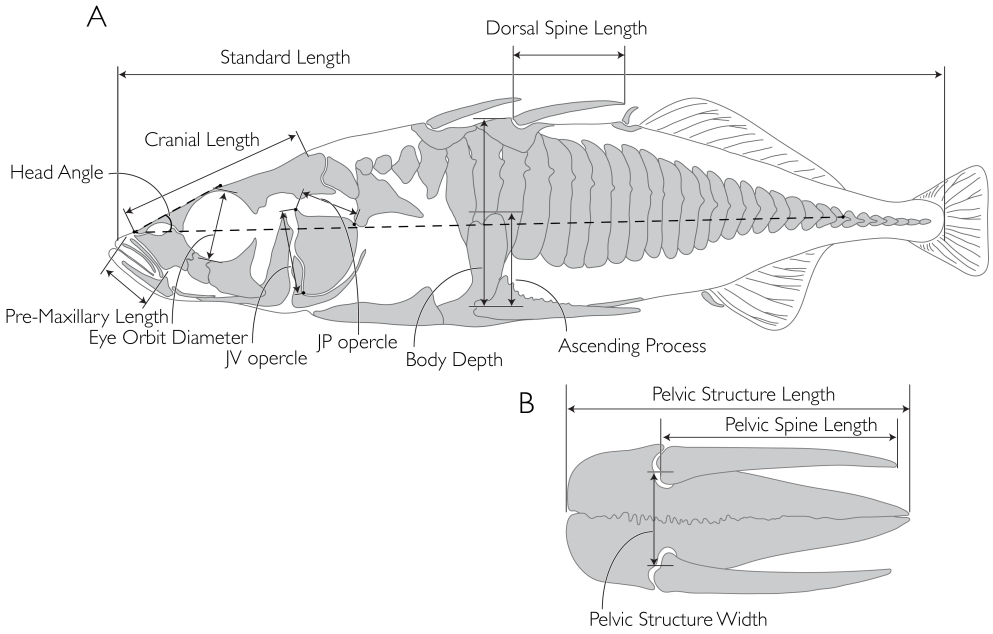

Figure 2.2. Linear and angular measures used for phenotypic analysis. (A) Lateral measurements. Including: pre-maxillary length, eye orbit diameter, JV opercle, JP opercle, cranial length, body depth, ascending process, pelvic spine length, standard length, and head angle (B) Linear measurements of the pelvic structure. Including: pelvic structure length and width, and pelvic spine length.

Morphological traits were log-transformed prior to statistical analysis. All of these traits increased linearly with standard length (all, $P < 0.001$, regression analysis). Size standardization of morphological traits was calculated according to equation 1:

$$\log y_{ij} = \log y_{ij} - \beta(\log x_i - \log \bar{x}) \quad [1]$$

Where $\log y_{ij}$ is the size standardized value for trait j in individual i, $\log y_{ij}$ is the original log transformed value, β is the slope of the regression between trait j and standard length (SL) using a model II regression, $\log x_i$ is the SL of individual i, and $\log \bar{x}$ is the mean SL of all of the individuals in the dataset. For most traits, the relationship

between the individual traits and SL differed among populations (most, $P < 0.001$ trait x SL interaction, ANCOVA) and population specific slopes, β, for each trait were used to size standardize. In some populations, a significant interaction between some traits and SL was not found. Finley Creek Swamp, Green Island, and Pudding Creek did not demonstrate a significant interaction between dorsal spine length and SL and Pudding Creek did not demonstrate a significant interaction between pelvic spine length and SL. Unstandardized log transformed scores were used for these trait scores for these populations in the subsequent analysis.

Additionally, apparent outliers visualized in regression analysis between trait *j* and SL were removed from the analysis. A total of seven fish from four populations were removed and one trait score from one fish was removed. Outlier status appeared to be due to the fish being very small (under 2.0 cm SL) or the trait in question being developmentally malformed.

Analysis of phenotypic variation

Univariate statistics

The number of lateral plates, dorsal and pelvic spine length, pelvic structure states, and gill raker numbers have prominently been used in previous studies of North American stickleback distributions to characterize divergence between oceanic and freshwater populations and divergence among freshwater populations. We use these characters to investigate traits in basin populations that demonstrate large differences in morphology as has been done in previous research (Hagen and Gilbertson, 1972; Reimchen et al., 1985; Bell and Foster, 1994; Bell and Ortí, 1994; Cresko et al., 2004;

Reimchen and Nosil, 2006). We do this to identify traits that may be locally adapted to particular freshwater habitats. Here we report and compare raw mean values of the meristic traits and mean values of size standardized morphometric traits of those mentioned above. These scores are reported as percentages of the mean standard length.

Principal component analysis

To visualize how individuals and populations were distributed in phenotypic space an ordination technique called Principal Component Analysis (PCA) was implemented. This analysis was performed using the R (R Development Core Team, 2013) package pcaMethods (Stacklies et al., 2007) with all 11 log transformed size standardized morphometric traits and all 3 meristic traits. Using this package all variables were scaled with unit variance and centered before ordination.

Approximately 2% of the data was missing due to absent or broken elements. A probabilistic PCA, which uses a maximum-likelihood approach, was used to infer missing values. Analysis was done including and excluding individuals with missing values to determine the robustness of the results. The percent variance explained for each principal component (PC) and average PC score for each population was robust between these analyses. The analysis that included individuals with missing values was slightly more conservative (results not shown).

Additionally, in some collections small fish were collected (2.0 - 3.0 cm SL). These small fish were probably young of the year. A sensitivity analysis was performed to determine if including these small individuals in the analysis drastically changed the inference made. PC analyses were performed on reduced datasets with individuals less

than 2.5, 3.0, 3.5, and 4.0 cm SL successively removed. Each of these reduced datasets were then compared to the original dataset, all 982 individuals, by regressing PC 1, PC 2, and trait loadings between these reduced datasets and the original values. All regressions were robust except when most or all individuals from a population dropped out of the dataset (results not shown). The results reported herein include analysis done with the full data set of 982 individuals and 29 populations that included individuals with missing values.

Principal component analysis excluding lateral plate counts

Stickleback with extra plates have been observed in the Willamette Basin prior to this study (Rutter, 1896). Extra lateral plate phenotypes in freshwater habitats could indicate individuals or populations that have anadromous or oceanic life histories that have been recently introduced through natural or artificial processes. However, if selection for reduced plates is weak, extra plated individuals can persist in freshwater habitats over long periods of time. Additionally, there is evidence that suggests that extra lateral plates can result as a secondary adaptation in freshwater populations (Kitano et al., 2008; Reimchen et al., 2013). If other genetically unlinked traits also segregate with oceanic populations it is a good indication that these individuals/populations have recent anadromous or oceanic life histories. To test this, an additional PC analysis was done excluding the lateral plate trait. Clustering of populations was visualized to determine if populations with extra lateral plate counts grouped with oceanic populations regardless of the presence or absence of extra lateral plates.

Analysis of the distribution of phenotypic variation

Analyses of Variance (ANOVAs) were performed using scores from PCs 1-3 to test the hypothesis that phenotypic variation is partitioned between coastal and Willamette Basin populations, between populations that inhabit the East and West sides of the basin, and among populations within the Willamette Basin. Extra lateral plate individuals were all found in the Eastern regions of the basin. To better estimate the phenotypic divergence between East and West regions free from potential influence of oceanic forms, additional ANOVAs were performed excluding populations that clustered with oceanic populations in the above analysis. Effect size, η^2, was calculated by dividing the sum of squares treatment by the sum of squares total from the ANOVA analysis.

Analysis of a phenotypic cline

The existence of phenotypic clines in nature can be caused by several different mechanisms. A phenotypic cline can be the result of localized adaptation to environmental heterogeneity or gradient, a result of drift, or it could be an indication of a hybrid zone in which two phenotypically divergent populations overlap and interbreed (Haldane, 1948; Endler, 1977; Streisfeld and Kohn, 2005). To investigate an apparent phenotypic cline in the McKenzie River, regression analyses were performed. The distance of each population from the mouth of the McKenzie was regressed against the average number of lateral plates and average PC 1 and PC 2 scores of each population.

RESULTS

The major axes of phenotypic variation distinguish oceanic and freshwater populations

The major axis of phenotypic variation, PC 1, explained 30% of the total variation and largely differentiates oceanic and Willamette Basin populations (Figure 2.3A). ANOVA confirmed that there was a significant difference in PC 1 scores between oceanic and Willamette Basin individuals ($F_{1,980} = 204.1$, $P < 0.001$, $\eta^2 = .172$; Figure 2.4A). The number of lateral plates, the number of gill rakers, body depth, pelvic structure width, orbit diameter, pre-maxillary length, and opercle JV & JP all have high loadings on PC 1 (Figure 2.3B, Table 2.2).

The second axis of phenotypic variation, PC 2, explained 24% of the phenotypic variation and also mostly differentiates oceanic and Willamette Basin populations. ANOVA confirmed that there was a significant difference of PC 2 scores between oceanic and Willamette Basin individuals ($F_{1,980} = 114.9$, $P < 0.001$, $\eta^2 = .105$; Figure 2.3A, 2.4A). Dorsal and pelvic spine length, ascending process length, pelvic structure length, dorsal cranium length, and pre-maxillary length all have high loadings on PC 2 (Figure 2.3B, Table 2.2).

The third axis of phenotypic variation, PC 3, explained 11% of the phenotypic variation and also significantly partitions phenotypic variation between oceanic and Willamette Basin individuals ($F_{1,980} = 235.2$, $P < 0.001$, $\eta^2 = .193$; Figure 2.4A). The ascending process, body depth, dorsal cranial length, and orbit diameter, and dorsal spine length all have high loadings on PC 3 (Table 2.2).

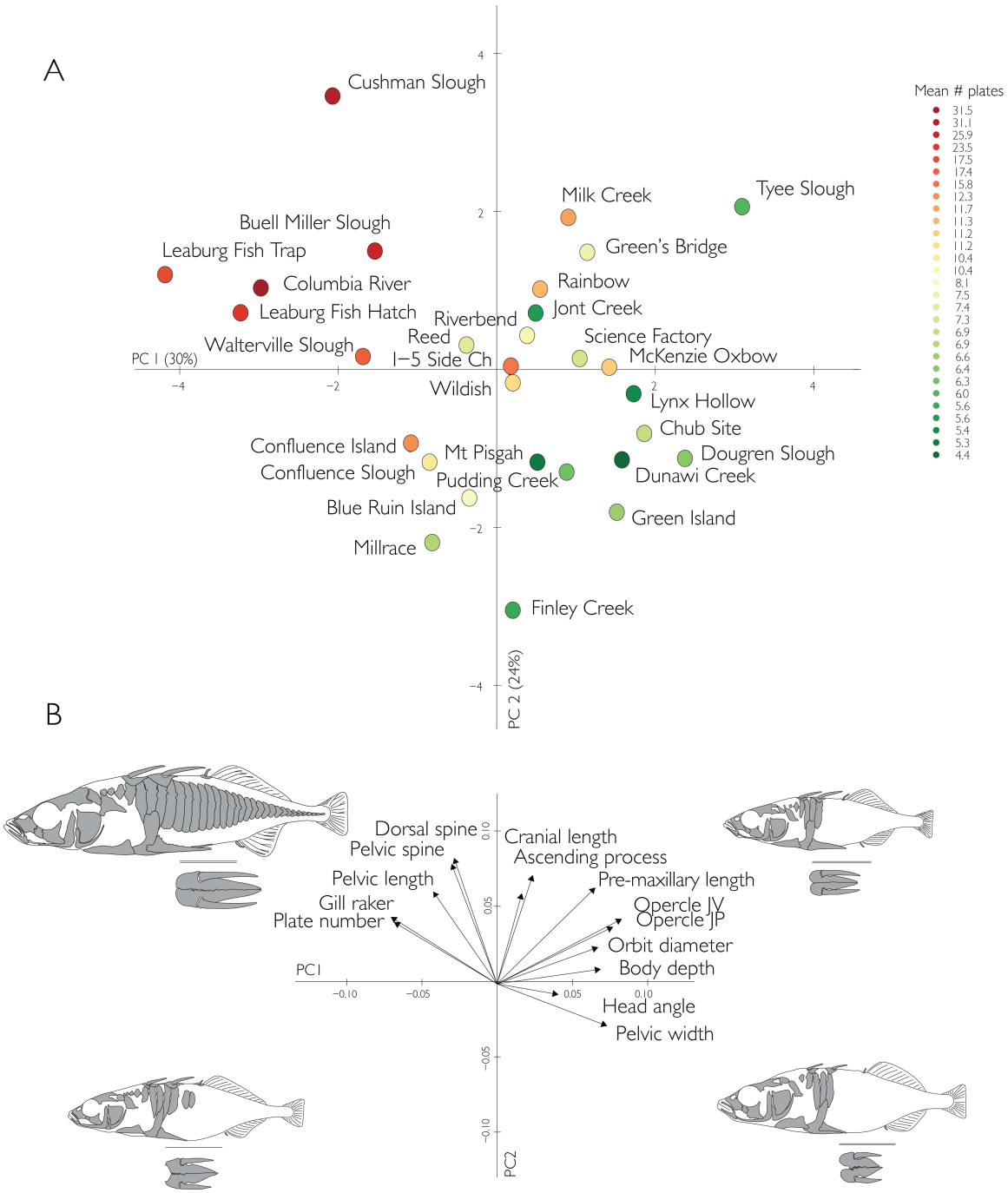

Figure 2.3. PCA with all 11 log transformed size standardized morphometric traits and all 3 meristic traits. (A) Distribution of average PC 1 and 2 scores for each population. PC 1 accounted for 30% of the total phenotypic variation and PC 2 accounted for 24% of the total phenotypic variation. Colors represent a heat map with red representing large average plate numbers and green indicating low average plate number. (B) Trait loadings along PC 1 and PC 2. Drawing of fish represent the extreme form of PC scores in the PC space where the drawing is located. Arrows indicate the direction of trait loadings and length indicated relative strength of that traits loading.

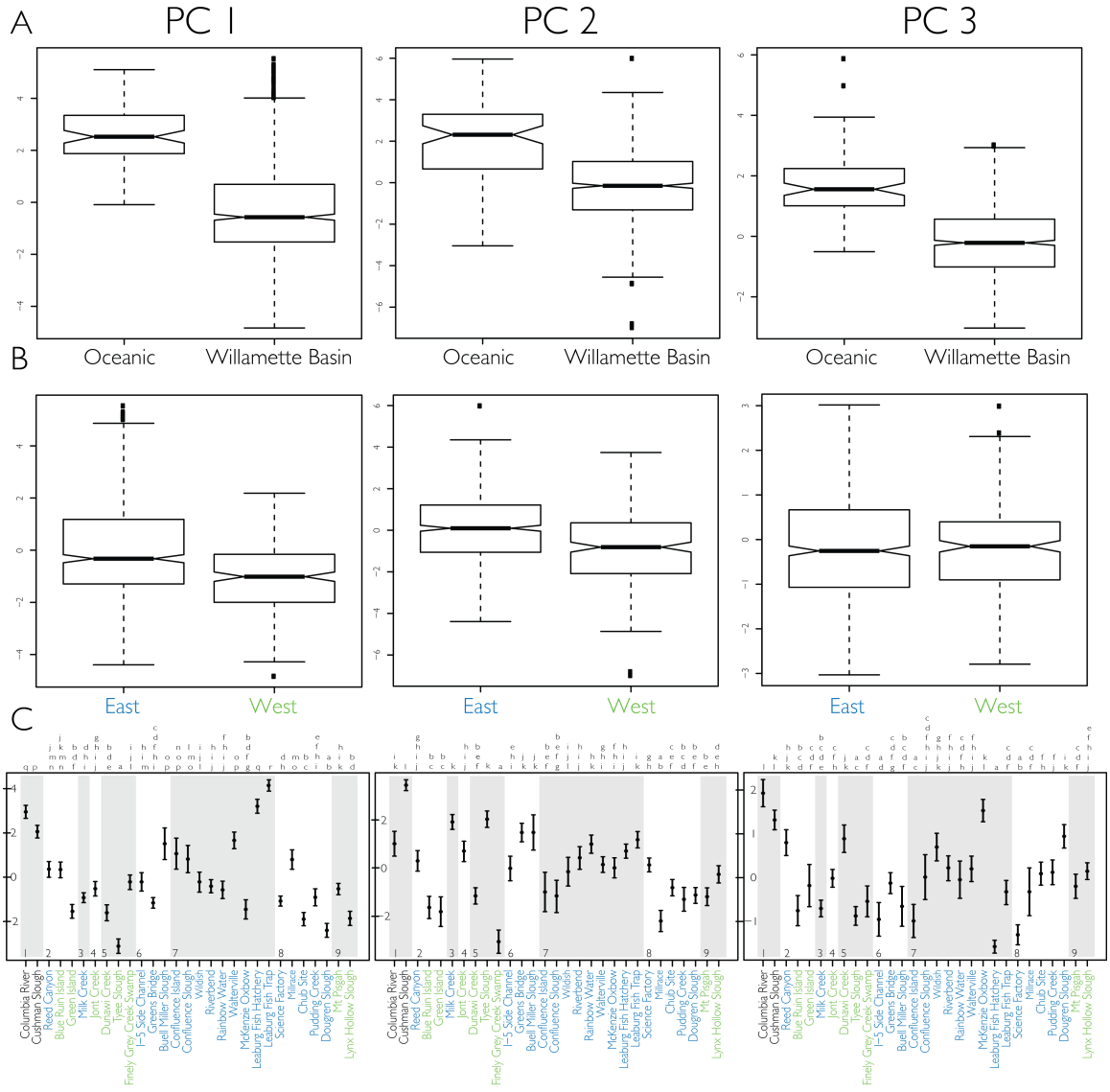

Figure 2.4. Box plots of regional and local comparisons for PCs 1-3. (A) Coast vs inland show significant phenotypic divergence for all three PCs. (B) Willamette Basin East vs West regional comparison. Slight but significant partitioning of phenotypic variation between the regions. (C) Individual population comparisons result in significant differences between populations designated by the lettering on top of the plots. Populations that share a letter are not significantly different. Populations colored in black are oceanic, blue are populations from the east, and green are populations from the west region. Grey banding is to signify different tributaries labeled with numbers bottom left of each band. 1- oceanic, 2 - Main Stem of the Willamette, 3 - Mollala, 4 - Luckiamute, 5 - Mary's, 6 - Santiam, 7 - McKenzie, 8 - Middle Fork, 9 - Coast Fork.

Table 2.2. Trait loadings of the first three PCs. Analysis done with all of the Oregon populations.

Trait	PC 1	PC 2	PC 3
Head Angle	0.191	-0.041	-0.102
Plate Number	-0.314	0.211	-0.188
Gill Raker Number	-0.322	0.228	-0.184
Dorsal Spine Length	-0.137	0.406	0.305
Ascending Process Length	0.077	0.306	0.392
Body Depth	0.317	0.048	0.336
Pelvic Spine Length	-0.129	0.426	0.236
Pelvic Structure Width	0.336	-0.144	0.287
Pelvic Structure Length	-0.194	0.314	0.027
Cranial Length	0.110	0.368	-0.445
Orbit Diameter	0.309	0.123	-0.446
Pre-Maxillary Length	0.301	0.325	0.039
Opercle JV	0.382	0.220	-0.127
Opercle JP	0.355	0.191	-0.086

There are phenotypically oceanic threespine stickleback in the Willamette Basin

Principal components analyses grouped populations from the Willamette Basin with oceanic populations (Figure 2.3A). In order to test whether the grouping of some inland fish with marine was robust to excluding a stereotypically marine phenotype (high lateral plate armor) we removed that trait and repeated the analysis. PCA excluding lateral plates resulted in the same populations grouping with oceanic populations along PC 1 and to a lesser extent along PC 2 (Figure 2.5, red and black dots grouping together). These populations include: Buell-Miller Slough from the Santiam and Walterville, Leaburg Fish Hatchery, and Leaburg Fish Trap from the McKenzie. We subsequently refer to these populations as phenotypically oceanic.

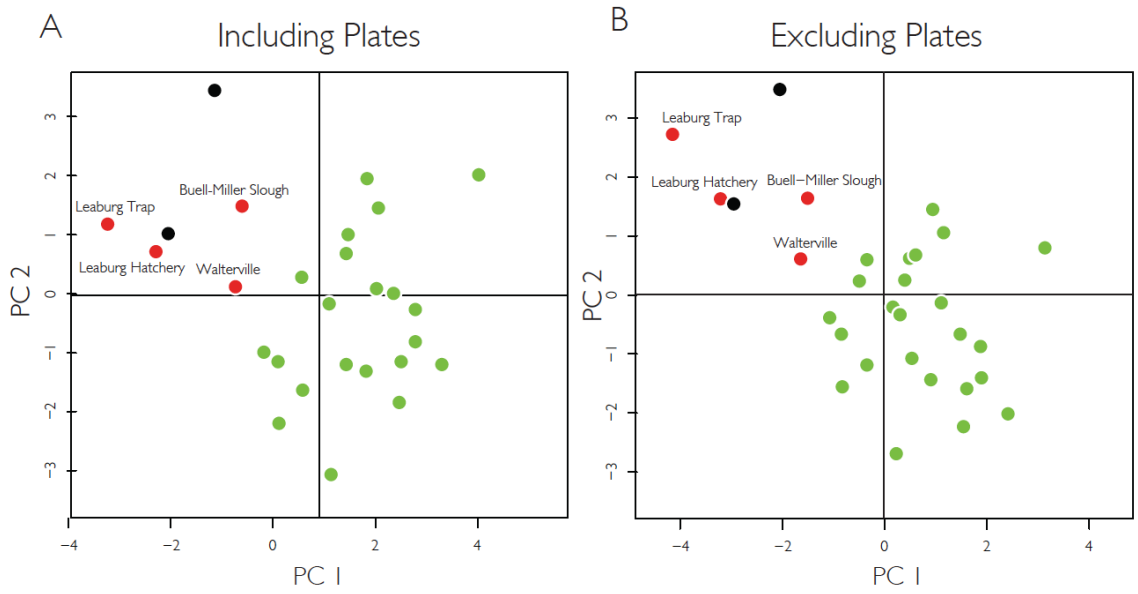

Figure 2.5. PC analysis including and excluding lateral plates used to identify populations that phenotypically group with oceanic populations regardless of the presence or absence of extra lateral plates. Green dots represents populations within the basin that are separated from the oceanic populations represented by black dots. The red dots represent the basin populations that cluster with oceanic populations in both analyses: Buell-Miller, Leaburg Fish Trap and Hatchery and Walterville Slough. (A) PCA including lateral plates, (B) PCA excluding lateral plates.

A phenotypic cline exists in the McKenzie River

Some phenotypic traits, most notably lateral plate number, appeared to become more ocean-like with increased distance from the confluence of the McKenzie-Willamette confluence. Linear regression between average plate number and river miles from the confluence resulted in a significant correlation ($F_{1,8} = 23.29$, $P = 0.001$, and $R^2 = 0.7124$; Figure 2.6). There were also significant associations between river miles and average PC 1 ($F_{1,8} = 8.03$, $P = 0.002$, and $R^2 = 0.4388$; Figure 2.6) and PC 2 scores ($F_{1,8} =$

12.32, $P = 0.008$, and $R^2 = 0.5571$; Figure 2.6) but no significant association between river miles and PC 3 scores ($F_{1,8} = 0.37$, $P = 0.560$; not shown).

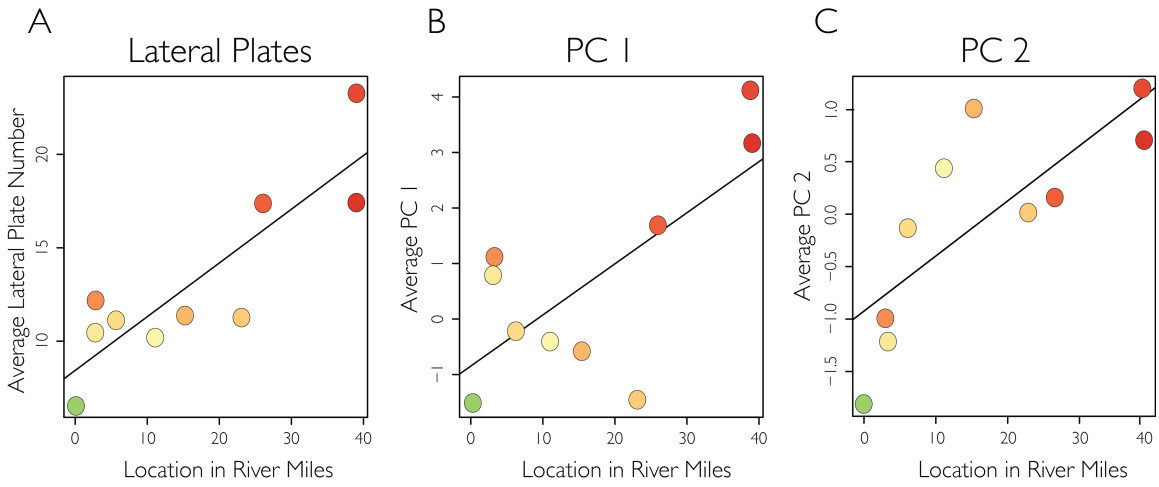

Figure 2.6. Regression analysis of the phenotypic cline in the McKenzie River. Regression were done between location of each population measured in distance from the McKenzie-Willamette confluence in river miles and (A) Average number of lateral plates, (B) Average PC 1 score, and (C) Average PC 2 score. All regressions resulted in significant correlations. Colors represent lateral plate counts as in figure 2.3.

There is little partitioning of phenotypic variation between the populations of the east and west regions of the Willamette Basin

Very little phenotypic variation was partitioned between populations that inhabit the east and west regions of the basin. There were slight but significant differences of PC 1 scores ($F_{1,889} = 69.21$, $P < 0.001$. $\eta^2 = 0.072$; Figure 2.4B) and PC 2 scores ($F_{1,889} = 29.3$, $P < 0.001$. $\eta^2 = 0.058$; Figure 2.4B) between populations that inhabit the East and West sides of the Willamette Basin, but both of these differences are very small as can be seen in the relatively small effect sizes. There was no significant difference in PC 3 scores ($F_{1,889} = 0.004$, $P < 0.95$. $\eta^2 = 0.00005$; Figure 2.4B).

The phenotypic partitioning that is present seems to be driven by the inclusion of particular populations in the East region. Buell-Miller, Leaburg Hatchery and Fish Trap,

and Walterville populations contained overtly phenotypically oceanic individuals, as determined by principal components analysis. When these populations were excluded from the analysis, there was still a slight but significant difference in PC 1 scores ($F_{1,738}$ = 11.43, $P < 0.001$. $\eta^2 = 0.015$) and PC 2 scores ($F_{1,889}$ = 54.75, $P < 0.001$. $\eta^2 = 0.038$), however the effect sizes became much smaller (data not shown).

There is significant partitioning of phenotypic variation among Willamette Basin populations

To investigate the overall patterns of phenotypic differences among basin populations, we analyzed the distributions of mean PC 1-3 scores of each population. Willamette Basin populations show significant differences in PCs 1-3 (Figure 2.3A, 2.4C). An omnibus ANOVA of each of these PCs resulted in significant differences between these populations in scores from PC 1 ($F_{26,864}$ = 90.16, $P < 0.001$), PC 2 ($F_{26,864}$ = 34.61, $P < 0.001$), and PC 3 ($F_{26,864}$ = 24.27, $P < 0.001$; Figure 2.4C). The only apparent pattern observed in the three PCs between or within individual drainages was that observed in the McKenzie River. Scores of PC 1-3 grew with distance from the confluence and trend toward oceanic PC scores (Figure 2.4C). One exception to this pattern is in PC 3 in which Leaburg Fish Hatchery, and to a lesser extent Leaburg Fish Trap, scores drastically decrease, opposing the trend seen in the other two PCs and away from oceanic. PC 1 scores for Tyee Slough are much lower than the other populations and are significantly different from all of the populations except Dougren Slough. Additionally, PC 2 scores are much lower and significantly different in Finley Grey Creek Swamp from all of the other populations except for the Mill Race (Figure 2.4C).

There are phenotypic trends that partition basin populations along axes that are orthogonal to the first two major axes of phenotypic variation (Figure 2.3A). The orthogonal axis that extends from - PC 1 and + PC 2 space to + PC 1 and - PC 2 space separates populations phenotypically in ways typical of oceanic to freshwater divergence seem on younger systems. Populations in -/+ space are phenotypically more oceanic-looking, characterized by high average plate counts, longer narrower pelvic structures, longer pelvic and dorsal spines, less sloped heads, and more gill rakers (Figure 2.3A, B). The perpendicular orthogonal axis separates populations phenotypically along the spectrum of freshwater morphologies. Populations in +/+ PC space are characterized by larger cranial facial features that include: jaw length, opercle JV and JP lengths, longer skulls, larger eye diameter. These populations also have a larger body depth than populations in -/- PC space (Figure 2.3A, B).

In general, Willamette Basin populations can be separated into low plated and extra plated morphologies. However, the presence of phenotypically oceanic populations in the basin suggests a recent introduction of oceanic fish that presents the possibility of introgression of oceanic genes into some basin populations. Below we report only those instances in which we identify differences in basin populations that might be free from recent oceanic influence; we define these as populations in which we encountered only low-plated fish.

There were large differences in spine and pelvic structure morphology in low plated populations in the basin (Table 2.3). The loss of dorsal or pelvic spines, or complete loss of the pelvic structure, as has been observed in Alaskan populations of stickleback (Bell and Ortí, 1994; Lescak et al., 2013), was not observed in any of the populations collected

in the basin. There was a very infrequent occurrence of a fourth dorsal spine in some basin populations (< 1%, data not shown).

Table 2.3. Means of traits used to characterize divergence between freshwater and oceanic and divergence between freshwater populations. Standard Length (SL) is in cm, plates and gill rakers are population averages, dorsal spine length, pelvic spine length, pelvic structure (PS) width and length, and orbit diameter are reported as % SL. Populations are arranged in order of average plate number. Populations in bold are the oceanic populations

Population	SL	Plates	Gill Raker	Dorsal Spine	Pelvic Spine	PS width	PS length	Orbit
Dunawi Creek	4.0	4.4	15.6	8.2	12.6	8.5	21.0	9.7
Mt Pisgah	4.1	5.3	13.2	8.9	14.3	8.4	21.4	8.5
Lynx Hollow Slough	3.5	5.4	14.4	11.0	16.1	11.2	23.2	10.5
Finely Grey Creek Swamp	4.1	5.6	13.9	7.9	12.6	10.2	20.3	8.5
Jont Creek	3.9	5.6	16.1	10.6	15.8	7.6	21.6	9.1
Tyee Slough	2.9	6.0	13.6	16.3	21.2	11.7	29.4	12.6
Pudding Creek	4.3	6.4	15.5	8.2	12.9	8.5	19.5	8.3
Dougren Slough	3.4	6.4	13.0	8.6	13.9	10.2	24.6	11.2
Green Island	4.0	6.6	15.6	8.3	12.6	10.0	18.8	9.0
Millrace	4.4	6.9	15.8	7.9	11.5	7.2	18.0	7.8
Chub Site	4.1	6.9	14.3	8.3	12.6	9.4	20.2	8.9
Blue Ruin Island	3.3	7.2	15.1	10.8	16.7	10.7	24.9	10.1
Science Factory	3.2	7.2	14.7	13.5	18.2	11.1	26.6	10.9
Reed Canyon	3.7	7.4	16.6	10.7	15.7	7.7	22.8	9.5
Greens Bridge	3.0	7.5	15.7	15.7	23.2	13.1	33.7	12.7
Confluence Slough	4.5	10.4	16.3	8.3	12.6	6.8	18.7	7.7
Riverbend	4.0	10.4	17.4	10.0	14.3	7.8	20.7	8.9
McKenzie Oxbow	3.8	11.0	16.6	9.2	14.6	9.4	21.3	10.2
Wildish	4.0	11.2	16.9	9.4	14.0	8.4	21.4	9.1
Milk Creek	3.6	11.7	16.2	12.5	19.3	9.5	27.3	10.1
Confluence Island	4.0	11.9	17.5	8.9	13.5	8.4	22.7	8.2
Rainbow Water	4.2	12.0	17.6	9.9	14.6	7.8	20.2	8.4
I-5 Side Channel	3.2	15.8	13.7	11.8	19.2	11.7	30.7	11.0
Walterville	5.2	17.4	19.5	7.7	11.3	5.4	17.2	6.5
Leaburg Fish Trap	4.8	17.5	21.5	9.5	16.3	5.0	21.0	6.6
Leaburg Fish Hatchery	4.3	23.5	19.0	11.3	15.4	6.4	20.8	6.9
Buell Miller Slough	4.1	25.9	17.1	11.4	17.9	7.9	24.3	8.6
Columbia River	**4.4**	**31.5**	**19.8**	**9.4**	**13.4**	**5.5**	**23.8**	**8.3**
Cushman Slough	**4.0**	**32.0**	**21.7**	**12.3**	**18.9**	**6.6**	**26.1**	**9.4**

Size corrected pelvic spine length was lowest in the Mill Race (~ 11% of SL) and highest in Green's Bridge and Tyee Slough (~ 22% of SL) and dorsal spine length was lowest in Mill Race (~ 8% of SL) and largest in Green's Bridge and Tyee Slough (~ 16% of SL). Pelvic structures (PS) in basin populations demonstrated large differences in

morphology. Pelvic length and width can vary independently. This is evidenced by some populations having short and narrow (PS) as in the Mill Race population (length ~ 18% of SL, width ~ 7% of SL), long and wide as in Green's Bridge (length ~34% of SL, width ~13% of SL), short and wide as in Green Island (length ~ 19% of SL, width ~ 10% of SL), or long and narrow as in Dougren Slough (length ~ 25% of SL, width ~ 8% of SL). Trophic morphology also demonstrated large differences in low plated basin populations. The average number of gill rakers ranged from ~ 13 in Dougren Slough to ~ 17 in Reed Canyon (Table 2.3). This is similar to what has been found in other freshwater stickleback (Hagen and Gilbertson, 1972; Reimchen and Nosil, 2006). Also the range in numbers of gill rakers is similar to differently diverged freshwater forms of stickleback (19-25) however much smaller in total number (Mcphail, 1984).

DISCUSSION

The phenotypic divergence in Willamette Basin populations from oceanic form is similar to what has been characterized in other freshwater populations

Phenotypic divergence from the oceanic form in older Willamette Basin populations is similar to what has been observed in younger freshwater populations. The first three principal components in this morphometric analysis explain 65% of the variation, and all three axes separate marine from Willamette Basin stickleback. The phenotypic departure of basin fish from marine is characterized by the loss of lateral plates, shortening of dorsal and pelvic spines, widening and shortening of pelvic structure, and reduction in the number of gill rakers, congruent with divergence described in younger <50 year old systems (Gelmond et al., 2009) and 15,000 year old systems

(Bell and Ortí, 1994; Bell, 2001; Cresko et al., 2004; Hohenlohe et al., 2010). There is growing evidence that much of the parallel phenotypic divergence between freshwater and oceanic stickleback populations from around the globe is the result of the continual reuse of standing genetic variation. This genetic variation has accumulated and is maintained by continual gene flow between a large panmictic marine population and countless independently founded freshwater populations (Colosimo et al., 2005; Barrett and Schluter, 2008; Hohenlohe et al., 2010). Our findings are consistent with the likelihood that this standing genetic variation predates the last ice age, and that Oregon freshwater genomes will share many features described in the genomes of geographically distant freshwater fish (Colosimo et al., 2005; Hohenlohe et al., 2010; Jones et al., 2012).

There are phenotypically oceanic stickleback in the Willamette Basin

Finding extra-plated individuals far from the ocean in the McKenzie, Santiam, and Mollala Rivers was unexpected. However, extra plated populations have been observed far inland in other Pacific Northwest (Rutter, 1896; Hagen and Gilbertson, 1972) and European stickleback systems (Münzing, 1963). The presence of these morphologies in the basin could indicate that selection has favored the maintenance or the reappearance of extra plates in certain basin habitats. Alternatively, high plated fish could be recently, perhaps artificially introduced, oceanic fish either by migration of anadromous forms into the basin or by anthropogenic transfer of fish from the coast into the basin.

Principal components analyses including and excluding lateral plate counts identified four populations that clustered closely with oceanic populations, suggesting

that multiple traits in these populations are similar to the oceanic form. These populations were Buell-Miller Slough from the Santiam River and Walterville Slough, Leaburg Fish Trap, and Leaburg Hatchery from the McKenzie River. Intensive fish collections at Leaburg Dam during the 1980's resulted in finding no stickleback, (Zakel and Reed, 1984). A public record of the presence or absence of stickleback in Buell-Miller Slough is lacking. The phenotypic evidence in the current study, however, along with the record of recent appearance of stickleback at Leaburg Lake strongly suggests that at least this population has been very recently introduced into the McKenzie River system from a coastal source, and the influence of this introduction is observable in downstream stickleback populations

There is a phenotypic cline in the McKenzie River that may indicate the presence of a hybrid zone

A phenotypic cline exists in the McKenzie River from the confluence with the Willamette to McKenzie river mile 39 as demonstrated by the correlations between river miles and average lateral plate number and phenotypic averages of PCs 1 and 2. It is feasible that this cline could be due to localized adaptation or drift (Haldane, 1948; Endler, 1977; Streisfeld and Kohn, 2005). However, in light of the apparent recent introduction of oceanic fish into Leaburg Lake it is possible that the observed cline is the result of introgressive hybridization between recently introduced oceanic fish and their downstream neighbors. The McKenzie Oxbow population (RM 22) deviates from the cline observed between PC 1 and river miles (Figure 2.4C, 2.6B). This population is located in a body of water that appears to have only intermittent connection with the main

stem of the McKenzie, and may be insulated from gene flow from Leaburg genotypes. The McKenzie River likely represents a hybrid zone between very young, phenotypically oceanic populations at Leaburg and older phenotypically freshwater populations lower in the drainage.

There is a lack of phenotypic divergence between stickleback populations that inhabit the east and west regions of the Willamette Basin

Despite the large differences in fluvial morphologies between the East and West regions of the basin we found very little phenotypic divergence between these groupings. The little divergence that was seen can be mostly attributed to the influence of the phenotypically oceanic populations in the McKenzie and the Santiam Rivers. The lack of divergence may be due to unexpected homogeneity of habitats between the two regions that stickleback inhabit. Stickleback were mainly collected in slow moving side channels and back waters and were not found in waters with large amounts of flow. Local habitats within and between each region inhabited by stickleback appear to be very similar and the differences in fluvial morphology that separate the East and West tributaries may not apply.

Phenotypic evidence of localized freshwater adaptation in defensive and trophic morphology

The only apparent pattern in individual population PC scores were those observed and accounted for by the cline in the McKenzie River. However there were some large

differences in individual PC scores in some populations. This suggests that some populations diverged significantly from other populations in certain traits.

PC axes 1 and 2 separated basin populations mostly along a stereotypic oceanic to freshwater transition. This was exemplified by the lateral plate phenotype. We did find populations with extra lateral plates in the basin. The existence of extra lateral plates has been hypothesized to be a post-capture adaptation in clear freshwater systems due to the increased frequency of capture by toothed predators because of the high clarity of the water (Reimchen et al., 2013). It is unclear however how other phenotypic traits segregate in these populations, and if these other traits trend towards oceanic like we have seen in the basin populations. The rivers in which we found extra plated individuals demonstrate high water clarity, which is consistent with the above adaptive hypothesis. Despite this, the clustering of the extra lateral plate with oceanic populations in analyses including and excluding the lateral plate phenotype and the clear evidence of a recent appearance of stickleback into one of the rivers in the basin suggests that the existence of individuals with extra lateral plates in the basin is due to a very recent introduction and not due to the processes of natural selection.

The other orthogonal axis separated populations along differences in freshwater morphologies. Dorsal and pelvic spine length and pelvic structure morphologies differed greatly in low plated populations. The axis described a vector between populations with long dorsal and pelvic spines and differing combinations of length and width of pelvic structures.

Spine lengths and pelvic morphologies have been previously shown to vary greatly in other freshwater distributions of stickleback (Hagen and Gilbertson, 1972; Bell and

Foster, 1994; Bell and Ortí, 1994; Webster et al., 2011; Millet et al., 2013). These traits are proposed to change in response to differences in calcium ions in the water and, in opposite directions, to the presence or absence of predatory fish or macroinvertebrates (Hoogland et al., 1956; Reimchen, 1983; Lescak et al., 2013). These findings suggest that dorsal and pelvic spine and pelvic structure morphologies may be under natural selection in the low plated Willamette Basin stickleback populations.

Trophic morphology also demonstrated large differences in some low plated basin populations. Gill raker numbers ranged from 13 to 17. This range in number of gill rakers is similar to what has been found in freshwater populations in Alaska and Washington State (Hagen and Gilbertson, 1972) and Haida Gwaii, BC (Reimchen and Nosil, 2006). The difference in the morphology is also similar to what has been observed in freshwater benthic and limnetic species pairs in British Columbia, (Mcphail, 1984) however these forms had much greater averages (19 to 25). Gill rakers are intimately associated with the types of food that stickleback forage upon and have been shown to be highly adaptable (Schluter, 1993; Day et al., 1994). This suggests that there is a range of food types in the different aquatic habitats in the basin and that stickleback populations are adapting to these food types. These findings support the hypothesis that freshwater populations of stickleback in the Willamette Basin have locally adapted to differing aquatic ecologies found throughout the basin.

CONCLUSIONS

We have shown that the stereotypical phenotypic change observed in young derived freshwater populations of threepsine stickleback likely originated before the last ice age.

This finding suggests an ancient origin of the standing genetic variation that appears to be continually reused during the oceanic to freshwater transition. Despite the lack of regional segregation of phenotypic variation in the basin we did find evidence of localized adaptation in spine, pelvic structure, and gill raker morphologies. These traits have been shown to be ecologically important and vary between freshwater forms in other systems. Future work is needed to determine if these traits are diverging adaptively. It will be interesting to understand if the repeated divergence that we observe between freshwater forms demonstrates a parallel genetic basis as appears to be the case in the oceanic to freshwater transition. Work is also needed to understand how genetic variation is partitioned in these populations and to compare this divergence with younger systems. We unexpectedly found the existence of phenotypically oceanic stickleback in the basin and the existence of a phenotypic cline in the McKenzie River. These fish are likely the result of a recent oceanic introduction and subsequent formation of a hybrid zone in the McKenzie River. Testing these hypotheses and investigating the genetic structure of Willamette Basin populations will require population genomic data.

BRIDGE

In chapter II we established that parallel phenotypic evolution in threespine stickleback predates the last ice age in the Willamette Basin of Oregon. We have also identified the existence of phenotypically oceanic populations of stickleback in the Willamette Basin and the existence of a phenotypic cline in the McKenzie River. We presented evidence of geographically localized phenotypic variation that may be indicative of adaptation in Willamette Basin populations.

In chapter III we investigate the regional and local distribution of genetic variation in Oregon focusing in on Willamette Basin populations of threespine stickleback. We also genetically test the hypothesis of a recent oceanic introduction in the basin and the hypothesis of a resulting hybrid zone in the McKenzie River. We also address questions of the patterns of genetic divergence through space and time.

CHAPTER III

THE GENETIC STRUCTURE OF THREESPINE STICKLEBACK (*GASTEROSTEUS ACULEATUS)* POPULATIONS FROM THE WILLAMETTE BASIN, OREGON

INTRODUCTION

Knowledge of how genetic variation is partitioned within and among populations is critical to inferring evolutionary mechanisms from patterns observed in nature (Dobzhansky, 1937; Dobzhansky, 1955; Wright, 1942; Wright, 1978; Haldane, 1955; Fisher, 1958; Endler, 1986). Research over the last century has led to a much better understanding of the evolutionary processes that mold and structure genetic variation through space and time (Wright, 1943a; Sokal et al., 1989; Hoekstra et al., 2004; Bustamante et al., 2005; Storz, 2005). However, until very recently these types of investigations were limited to a handful of genetic markers in very few individuals and very few populations (Allendorf et al., 2010; Etter et al., 2011; Hohenlohe et al., 2012). This restricted view of the genome limited the extent and generality of the inferences that could be made about evolutionary processes such as migration, genetic drift, and adaptation (Emerson et al., 2010; Nosil and Feder, 2012).

We are in the age of the genome. With recent advances in sequencing technology (Baird et al., 2008; Shendure and Ji, 2008; Davey et al., 2011; Glenn, 2011) and computational analysis (Pritchard et al., 2000; Meirmans and Van, 2004; Catchen et al., 2011) we now have a greatly expanded view of the genome that allows us to investigate the partitioning of genetic variation using thousands of markers spread throughout the

genome in many individuals and in many populations. This allows us to make detailed inferences of divergence over multiple temporal scales, make comparisons of homologous loci across populations, be able to differentiate neutral processes that randomly affect the entire genome from adaptation that change allele frequencies differentially in a genomically localized manner, and compare genome-wide patterns created by neutral and adaptive process between spatial and temporally distributed populations.

Contemporary investigations explore and have brought to light how evolutionary processes of migration, mutation, drift, and natural selection partition genetic variation within and between populations at the genome-wide level (Caicedo et al., 2007; Bryc et al., 2010; Vonholdt et al., 2010; Byrne et al., 2013; Catchen et al., 2013). Despite these advances it is still unclear how neutral divergence and adaptive divergence compare over different temporal and spatial scales.

The threespine stickleback (*Gasterosteus aculeatus)* fish is an exceptional species for the study of population genetic structure in vertebrates. Its relatively small (~ 460 Mb) genome has been sequenced and made available for reference (Jones et al., 2012). Much work has been done to understand how populations of stickleback genetically diverge and we have found genetic regions and possibly genes important for adaptive traits (Shapiro et al., 2004; Colosimo et al., 2005; Hohenlohe et al., 2010; Jones et al., 2012). Surprisingly, it appears that the same genes and genomic regions are continually reused during adaptation from ocean to freshwater (Colosimo et al., 2005; Hohenlohe et al., 2010; Jones et al., 2012). Despite the recent advances of population genomics in threespine stickleback populations, a gap in our knowledge stems from the fact that much

of this work has been restricted to relatively young (~13,000 years old) populations. These populations result from invasion of marine populations into freshwater habitats that formed after glacial recession and subsequent continental rebound at the end of the last ice age (Bell and Foster, 1994). Research is currently being done on very young populations (<50 years old) that resulted from the uplift of Middleton Island during the earthquake of 1964 (Gelmond et al., 2009). Little is known about stickleback populations that are much older. Is the magnitude of genetic divergence much greater in older systems than what has been seen in younger systems? How does population structure compare over different spatial scales?

Less work has been done on the numerous populations of stickleback that exist in aquatic habitats in lower latitudes that were not covered by glaciers, making their origins potentially much older than the last ice age. Many aquatic habitats in Oregon and those in the Willamette Basin are potentially millions of years old (Benner and Sedell, 1997; O'Connor, 2001; Boothl et al., 2003) and the stickleback populations that inhabit these ecosystems may therefore also be very old. This presents us with a unique system to investigate the genetic and genomic basis of divergence over multiple temporal and spatial scales. Previous research found a high degree of genetic divergence between Oregon oceanic fish and a Willamette basin population, much greater than what had been observed in younger freshwater populations (Catchen et al., 2013). However, this study sampled only one location so that the structure of genetic variation within this vast inland riverine system and correlations with environmental variables and genetic structuring was unable to be investigated.

In the current study, we used RADseq to investigate geographically broad and regional patterns of genetic structuring in a previously unexplored distribution of populations across the Willamette Basin. We tested the hypothesis of a recent introduction of an oceanic population in the Willamette Basin and the hypothesis of a hybrid zone in the McKenzie River that resulted from this introduction. Additionally, this work allowed us to address the questions of genetic divergence over different temporal and spatial scales.

METHODS

Collection of stickleback samples

Stickleback fish were collected for genotypic analysis as part of a larger set of collections used for phenotypic analysis (Figure 2.1, Table 2.1; Currey et al. 2014 (in prep)). Briefly, collections were made using minnow traps as previously described (Cresko et al., 2004; Catchen et al., 2013). Up to 100 stickleback from each collecting location were euthanized using MS222 and fixed in 95% ETOH. GPS coordinates of each location were obtained using Google Earth. Samples were also obtained from state and federal agencies, namely Oregon Department of Fish and Wildlife (ODFW) and National Oceanic and Atmospheric Administration (NOAA), which were conducting collections of other fish species in the Willamette and Columbia basins.

Populations for genetic analysis were chosen with three goals in mind. First, to investigate genetic structure of stickleback in the Willamette Basin, populations were selected from most of the major tributaries in the basin. Second, additional populations were selected from along the McKenzie River to obtain a fine scale view of the genetic

structure in a subset of basin stickleback. These populations also allowed us to investigate the genetic structure of the observed phenotypic cline as described in previous research (Currey et al. 2014 (in prep)). Lastly, one additional population, presumably oceanic, from the mouth of the Columbia was selected. Eight previously investigated Oregon populations (Catchen et al., 2013) five located along the Oregon coast, one (previously mentioned) from the Willamette Basin, and three from the southern end of the Deschutes drainage (Central Oregon) were included, thus providing wide regional coverage of stickleback populations (Currey et al. 2014 (in prep)). These efforts provided a total of 1031 individuals in 20 populations that were used in this analysis.

DNA extraction

A Qiagen DNeasy Blood and Tissue Kit was used to extract DNA. Quantification of DNA concentration was performed using a Qubit fluorometer using the dsDNA broad range assay kit. DNA dilution plates for each population were created by diluting to a concentration of 25 ng/ul using EB in 96 deep well format.

RAD sequencing and SNP discovery

RAD libraries were created as in Catchen et al, 2013. Briefly, for each population sequenced, endonuclease *SbfI-HF* (New England Biolabs) was used to cut DNA at specific locations spread throughout the genome. To the digested DNA of each individual within each population P1 adaptors with unique barcodes were ligated so that sequences from each individual could be bioinformatically recovered. Individuals were then pooled and sheared. Sheared fragments of 300 - 500 bp in length were size selected and a second

adapter, P2, was ligated to the fragments. Primers specific to sequences located on the P1 and P2 adaptors were used to perform a PCR reaction to enrich only for fragments with both a P1 and P2 adaptor in 12 cycles of amplification. The resulting amplified library was size selected by gel electrophoresis for fragments of 388 - 520 bps and then these fragments were sequenced using an Illumina HiSeq2500 at the University of Oregon Genomics Core Facility.

The *Stacks* software pipeline (Catchen et al., 2011) and `process_radtags` was utilized to process the resulting RAD reads. This processing included de-multiplexing the libraries by barcode and filtering to a desired sequencing quality score (phred score). Reads with a quality score below 90% probability of being correct were eliminated from the rest of the processing. Processed tags were then aligned against the stickleback genome using GSNAP (Wu and Nacu, 2010). Aligned tags were then run through the Stacks pipeline using `rxstacks` to identify and call SNPs throughout the genome and create a catalog of SNPs from all populations included.

Calculation of population genetic statistics

The `populations` program within the *Stacks* framework was utilized to calculate population genetic measures of F_{ST}, π, and F_{IS} for each SNP. Genome-wide estimates of π and F_{IS} were computed by averaging over all variable sites and over all variable and non-variable sites. To investigate the genetic relatedness between Willamette Basin and oceanic populations and among Willamette Basin populations we calculated an average genome-wide measure of F_{ST} for all pairwise combinations among and between these two groupings.

Thousands of loci are generated from this method and are not computationally manageable for most software packages. To create a manageable subset of loci, 1000 randomly sampled SNPs were generated using a Perl script as described in the *Stacks* manual. Populations were subdivided to analyze genetic relatedness of the populations in that subgroup (e.g. only Willamette Basin populations). For each subgroup analysis a new set of 1000 loci was generated using *Stacks* by running only the populations included in each particular subset through the `populations` program.

Population genomic analyses

To visualize population structure and substructure multiple groupings were investigated using principal component analyses (PCA) and STRUCTURE analysis (Pritchard et al., 2000; Falush et al., 2003). Two broad-scale analyses were performed, one with all of the Oregon populations and another with all of the Willamette Basin populations. The state of lateral plate morphology has been classically and consistently used in stickleback research to group individuals and populations into different phenotypic and life history classes (Hagen and Gilbertson, 1972; Reimchen and Nosil, 2006). Previous phenotypic analysis (Currey et al. 2014 (in prep)) identified two groupings of populations in the basin, one consisting of populations with the low plated phenotype and another with populations consisting of various degrees of the extra plated phenotype. Willamette Basin populations were divided into these two plate classes, low and extra, for further substructure analysis.

Principal component analysis of genetic variation

To investigate and visualize the axis of genetic variation, a PCA was performed on each subset of 1000 randomly chosen loci corresponding to the subgroups defined above using the software Genodive (Meirmans and Van, 2004). The mean and standard deviation of PC1 and PC2 were analyzed using R (Team, 2012). For all PCAs, a covariance matrix was used and significance was tested using a resampling method with 1000 permutations.

Population structure analysis

Population structure was analyzed using the same 1000 randomly sampled loci for each population subset used in the individual PCA analyses above. STRUCTURE analyses were performed on each grouping using a burn-in of 20,000 steps, and with 20,000 replicates for each value of *K*, where *K* is the number of groupings. The number of *K*s tested for each subset was equal to the number of populations in each grouping. Each value of k was replicated 10 times and the Evanno method (Evanno et al., 2005) using Harvester (Earl, 2012) visual inspection of the change in *Ln P(D)*, and biological relevance was used to determine the level of *K* that best fit the data.

Analysis of molecular variance

The partitioning of genetic variation within and among populations was determined using an Analysis of Molecular Variance (AMOVA) (Excoffier et al., 1992) that was performed using the software Genodive. We hypothesized the existence of genetic partitioning between coastal and inland populations, between extra and low plated basin

populations, and among populations within the extra and low plated groupings. We therefore performed our AMOVA analyses to test these hypotheses. All of these expectations were tested using the 1000 randomly sampled loci specific to the subset of populations unique to each analysis. Coastal and Willamette Basin populations were designated by regional location and not by phenotype. All AMOVAs were done using an infinite allele model. Significance was determined by resampling using 1000 permutations.

Analysis of genetic structure within populations

Wright's inbreeding coefficient, F_{IS}, was used to look for structuring of genetic variation among individuals within each of the basin populations (Wright, 1942). Values of F_{IS} can theoretically range from -1 to +1, with negative values indicating an excess of heterozygotes and positive numbers indicating an excess of homozygotes. F_{IS} values that deviate positively from Hardy-Weinberg Equilibrium can be an indication of assortative mating or cryptic genetic structure (Nei, 1975; Hartl and Clark, 1997).

Analysis of isolation by distance

Populations that are physically closer to one another are expected to be more genetically similar than populations that are farther apart because of the consequences of limited dispersal, a pattern termed Isolation by Distance (IBD) (Wright, 1943b). IBD theory predicts genetic divergence should scale with physical distance so that there should be a high degree of correlation between geographical distance and genetic divergence between populations. To identify signals of IBD within Willamette Basin

populations we used a Mantel's test to calculate correlations between genetic distance, using averaged genome wide measures of F_{ST} and a geographical distance measure. River systems are very dynamic, and we therefore decided to include river gradient when testing hypotheses of IBD. To incorporate river gradient in the measure of geographic distance we employed the method proposed by Castric *et al*, 2001 (Castric et al., 2001). In this method elevation is used as a proxy for river gradient assuming that differences in elevation will correlate strongly with river gradient. The sum of difference in elevation (SOE) was calculated using the equation 2:

$$SOE = (e1 - eN) + (e2 - eN) \quad [2]$$

Where SOE is the sum of elevation differences, e1 and e2 is the elevation of any pair of populations, and eN is the elevation of the first common river node. Mantel's tests were performed using the software Genodive and significance was found using a resampling method with 10000 permutations. The apparent presence of oceanic populations within the basin may confound inference made with this analysis so a second analysis was done excluding the phenotypically oceanic populations identified in the phenotypic analysis (Buell-Miller, Walterville Slough, Leaburg; Currey et al. 2014 (in prep)).

Analysis of a genetic cline in the McKenzie River

Regression analysis was used to look for the correlations between the distance in river miles from the McKenzie-Willamette confluence and average genetic PC 1 scores of populations in the McKenzie River. An additional regression was done between the

same average genetic PC 1 scores and average phenotypic PC 1 scores to investigate the correlation between the observed phenotypic and genetic clines in the McKenzie River.

RESULTS

Genetic variation in Oregon populations of stickleback is partitioned between coastal and inland populations

Analysis of the genetic variation in all twenty populations differentiates coastal and inland populations in both PC and STRUCTURE analysis. This is consistent with what was found in a previous study using a subset of this data (Catchen et al., 2013). The first PC, which accounts for ~ 62% of the total genetic variation and is the only significant PC (Table 3.1) separates coastal from inland populations (Figure 3.1A). STRUCTURE analysis produces the same result. Using *Ln P(D)* and the deltaK method (Evanno et al., 2005) we found that a model with K =2 best fits this full data set. The plot of posterior probabilities for this level of K recapitulates the two groupings found in the PC analysis (Figure 3.1B). We used AMOVA to investigate how genetic variation is more precisely partitioned within and between individuals, populations, and regions (coastal and inland). We found that most of the variation not attributed to within individuals (~58%) was partitioned between the two geographical regions (~22%) with a lesser amount partitioned between populations (~16%; Table 3.2).

Average measures of population differentiation demonstrate large amounts of divergence between oceanic and Willamette Basin populations. On average, F_{ST} values between oceanic and Willamette Basin low-plated populations was ~ 0.13, with a range of 0.10 (Cushman Slough vs Finley Grey Creek Swamp) to 0.17 (Millport Slough vs

Dougren Slough). This is similar to the amount of divergence that was found between oceanic and a single basin population in a previous study (Catchen et al., 2013). Interestingly, the average amount of divergence between low plated basin populations and oceanic populations decreases with an increase in distance (measured as the fish swims). In other words, average F_{ST} between low plated basin and coastal populations is greater with respect to northern coastal sites (Columbia and Millport Slough) than more southern sites (Cushman Slough and South Jetty), respectively: 0.138 and 0.155 versus 0.10 and 0.11 (Table 3.3).

Table 3.1. PCA results of analysis done with all 20 Oregon populations. Eigenvalues, percent variance explained, and p-values of PCs of genetic variation. The first PC accounts for the vast majority of the genetic variation and is the only axis that is statistically significantly.

	Eigenvalue	% Variance	Cumulative	*P*-value
1	13.172	57.33	57.33	0.01
2	2.536	11.038	68.368	0.87
3	1.441	6.271	74.639	1
4	1.097	4.775	79.414	1
5	0.879	3.824	83.238	1
6	0.781	3.401	86.639	1

Phenotypically oceanic populations within the Willamette Basin genetically group with coastal populations

PC and STRUCTURE analyses demonstrate an exception to the coastal and inland groupings. Leaburg, and to a lesser extent Buell-Miller and Walterville Sloughs, group with coastal (and oceanic) populations along PC 1 (Figure 3.1A) and share a majority of their ancestry with the coastal grouping in the STRUCTURE analysis at $K = 2$, the model that best fits the data (Figure 3.1B). Tellingly, these are the same populations that

grouped with the oceanic populations in the phenotypic analysis (Currey et al. 2014 (in prep)). All of the populations that have extra plates in the basin appear to share ancestry with both coastal and inland populations (Figure 3.1B, indicated in red).

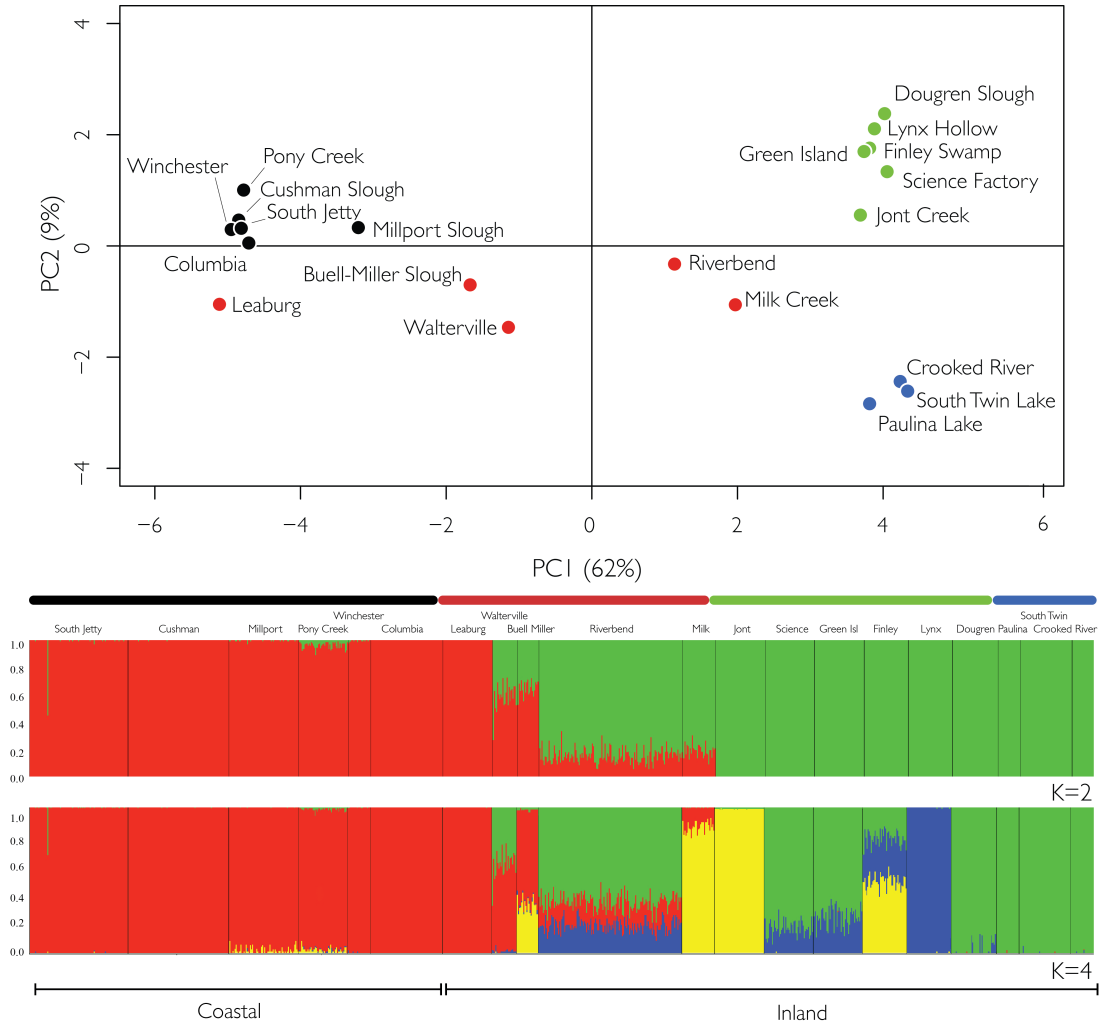

Figure 3.1. PCA and STRUCTURE plots of all of the Oregon populations in this analysis. (A) PC 1 explains 62% of the genetic variation and separates coastal and inland populations. Dots represent the average PC 1 and PC 2 score of each population. Red dots represent Willamette Basin high plated populations, Blue represents Deschutes Basin populations, Green dots represent Willamette Basin low plated populations, and Black represents coastal populations. (B & C) Plots of posterior probabilities for cluster assignment for $K = 2$ (B) and $K = 4$ (C). $K = 2$ was found to be the model that best fit the data. Each vertical bar represents a single individual within a population (separated by black vertical bars and labeled above). The proportions of the color in each column represent the posterior probability of cluster assignment for that individual. Colored bars above the plot represent the same populations groupings explained for the dots.

Table 3.2. AMOVA table demonstrating the amount of partitioning of genetic variation between Coastal and Inland groupings. This analysis clearly shows that the majority of variation not accounted for by individuals or populations is partitioned between Coastal and Inland populations.

Source of Variation	Nested in	% Variance	*F*-stat	*P*-value
Within Individual	----	0.578	F_{IT}	----
Among Individual	Population	0.048	F_{IS}	0.001
Among Population	Coastal vs Inland	0.156	F_{SC}	0.001
Among Coastal vs Inland	----	0.217	F_{CT}	0.001

Table 3.3. Genome-wide averaged F_{ST} pairwise comparisons for all of the Willamette Basin and oceanic populations. D = Dougren Slough, LH = Lynx Hollow, FS = Finley Swamp, GI = Green Island, SF = Science Factory, JC = Jont Creek, MC = Milk Creek, R = Riverbend, BM = Buell-Miller, W = Walterville, LB = Leaburg Dam, C = Columbia River, MS = Millport Slough, CS = Cushman Slough, and SJ = South Jetty.

	LH	FS	GI	SF	JC	MC	R	BM	W	LD	C	MS	CS	SJ
D	0.431	0.165	0.128	0.090	0.309	0.245	0.069	0.201	0.279	0.458	0.147	0.165	0.112	0.117
LH		0.181	0.144	0.142	0.311	0.244	0.073	0.199	0.277	0.462	0.148	0.164	0.113	0.118
FS			0.072	0.075	0.149	0.155	0.051	0.153	0.211	0.370	0.130	0.145	0.103	0.106
GI				0.070	0.157	0.163	0.046	0.161	0.222	0.373	0.134	0.149	0.105	0.109
SF					0.161	0.155	0.049	0.153	0.212	0.370	0.131	0.147	0.104	0.107
JC						0.196	0.074	0.178	0.248	0.422	0.141	0.158	0.110	0.114
MC							0.064	0.100	0.159	0.321	0.105	0.122	0.085	0.087
R								0.051	0.043	0.155	0.064	0.080	0.067	0.064
BM									0.047	0.125	0.039	0.052	0.034	0.035
W										0.113	0.045	0.056	0.036	0.037
LD											0.041	0.046	0.026	0.027
C												0.009	0.005	0.006
MS													0.008	0.009
CS														0.001

F_{ST} analysis also reveals similarity of these populations to coastal populations. The extra plated populations in the basin are much more closely related to oceanic fish than to low plated populations. F_{ST} between extra plated and oceanic (average is ~ 0.06) is significantly less than F_{ST} between low and extra plated basin populations (average is ~ 0.22; ANOVA, $F_{1,56} = 53.36$, $P < 0.001$). Interestingly, the average divergence between the phenotypically oceanic populations (Leaburg, Walterville, and Buell-Miller) and

oceanic is ~ 0.04 (compared to ~0.28 from the low plated populations). This is very similar to what has been observed between very young freshwater populations on Middleton Island and oceanic (average is ~ 0.03; (Lescak et al. 2014 (in prep)) suggesting that these populations have been very recently introduced.

A genetic cline exists in the McKenzie River

Genetic analysis supports the existence of a phenotypic cline in the McKenzie River. As reported previously, populations in the McKenzie River progressively become more phenotypically oceanic with river miles from the McKenzie-Willamette confluence (Currey et al. 2014 (in prep)). STRUCTURE analysis demonstrates the existence of a genetic cline in the McKenzie River with the posterior probabilities of the individuals in these populations progressively becoming more coastal (oceanic) with distance from the confluence (Figure 3.1B, 3.2B). Regression analysis between the average genetic PC 1 scores and river miles from the confluence demonstrates that there is a very strong correlation ($R^2 = 0.97$, $F_{1,2} = 113.1$, $P = 0.009$; Figure 3.3A). Additionally, there is a nearly perfect correlation between average genetic and phenotypic PC 1 scores in the McKenzie River populations ($R^2 = 0.98$, $F_{1,2} = 166.7$, $P = 0.006$; Figure 3.3B). F_{ST} analysis mirrors this same cline. The average F_{ST} of combined oceanic populations versus proximate to distal populations, Green Island, Riverbend, Walterville, and Leaburg, are 0.123, 0.069, 0.044, and 0.035, respectively (Table 3.3).

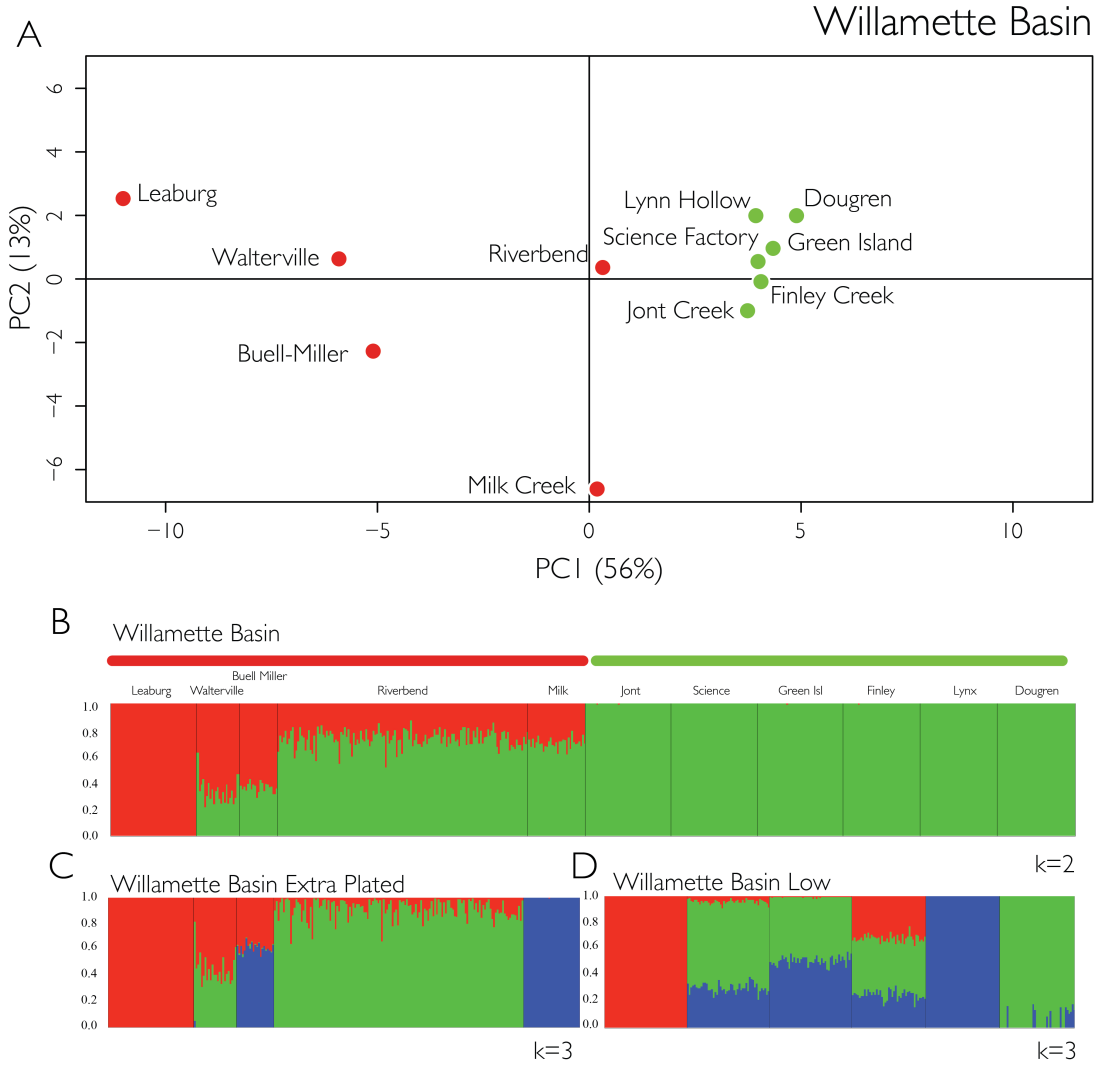

Figure 3.2. PCA and STRUCTURE plots of the Willamette Basin populations. (A) PC 1 explains 56% of the genetic variation and separates extra and low plated populations. Dots represent the average PC 1 and PC 2 score of each population. Red dots represent Willamette Basin high plated populations and green dots represents Willamette Basin low plated populations. (B) STRUCTURE analysis of all Willamette Basin populations. (C) STRUCTURE analysis of Willamette Basin extra plated populations. (D) STRUCTURE analysis of Willamette Basin low plated populations. (B, C, & D) Plots of posterior probabilities for cluster assignment for K = 2 (B) and K = 3 (C & D). All values of K represent the value of the model that best fits the data. Each vertical bar represents a single individual within a population (separated by black vertical bars and labeled above). The proportions of the color in each column represent the posterior probability of cluster assignment for that individual. Colored bars above the plot represent the same populations groupings explained for the dots.

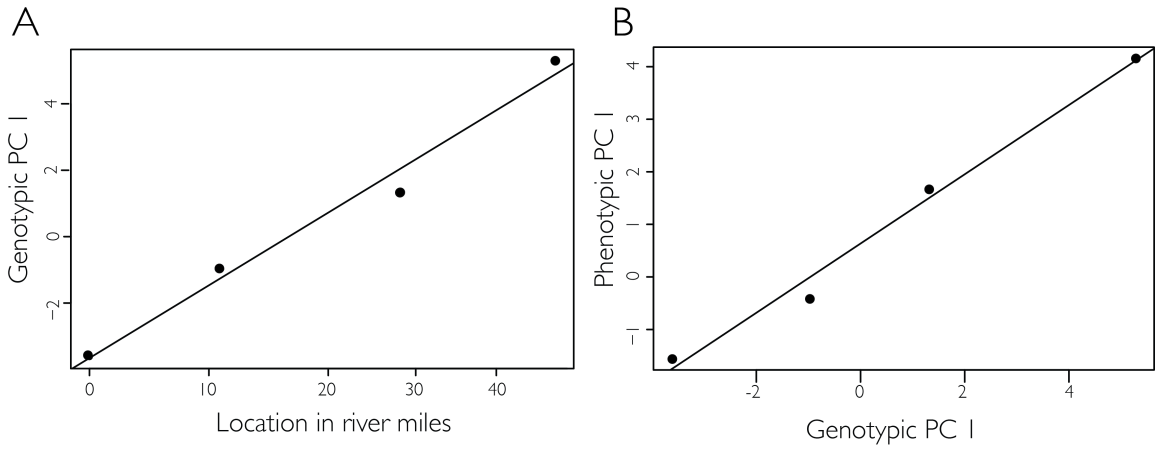

Figure 3.3. Regression analysis of a genetic cline in the McKenzie River (A) Location of each population from the Willamette-McKenzie confluence in river miles and average genotypic PC 1 score, $R^2 = 0.97$, and (B) average genotypic PC 1 score average phenotypic PC 1 score $R^2 = 0.98$.

Low plated populations had lower values of F_{IS} (~ 0.0001; all positions) compared to extra plated populations (~ 0.0003). It is interesting to note that the values of F_{IS} are relatively elevated in Riverbend, Walterville Slough, and Buell-Miller Slough respectively: 0.0005, 0.0002, 0.0005 (Table 3.4) suggesting cryptic genetic structure that would be expected in populations within a hybrid zone.

Structure of stickleback populations in the Willamette Basin

PCA differentiated populations along the PC 1 axis into low and extra plated groupings (Figure 3.2A). This PC accounts for ~ 56% of the total genetic variation and is the only significant PC (Table 3.5). We found that a model of K =2 best supports the data in a STRUCTURE analysis. A plot of this level of K shows these same groupings of low and extra plated populations (Figure 3.2B). By AMOVA, we found that a significant amount of variation was partitioned between extra and low plated populations (~13%)

that was not partitioned within individuals nested within populations (~54%) or partitioned between populations nested within low and extra plated groupings (~28%; Table 3.6).

Table 3.4. Genome-wide averaged F_{IS} and π statistics for all 20 populations for all nucleotides contained within a RAD site (variant and fixed). N is the average number of individuals sequenced, Private is the number of variable sites unique to each population, π is nucleotide diversity, and F_{IS} is Wright's inbreeding coefficient.

All positions (variant and fixed)				
Population ID	Private	N	π	F_{IS}
Dougren Slough	22	43.2493	0.000111556	0.000138036
Lynx Hollow	52	42.7947	0.00024392	6.67E-05
Finley Swamp	271	42.9061	0.000751442	5.62E-05
Green Island	240	47.804	0.000585515	0.000182351
Science Factory	103	47.9011	0.000760442	0.000155714
Jont Creek	208	44.9428	0.000476098	2.44E-05
Milk Creek	43	31.9183	0.00136512	0.00015006
Riverbend	289	136.477	0.00199151	0.000546338
Buell Miller	128	20.9494	0.00263037	0.000275323
Walterville	31	23.8923	0.00252663	0.000509787
Leaburg Dam	89	46.7043	0.00151578	0.000104798
Columbia	1345	47.9293	0.00216786	0.000800943
Millport Slough	924	66.5964	0.0019875	0.00114333
Cushman Slough	1895	96.4928	0.00226263	0.00118357
South Jetty	1523	86.1928	0.00235918	0.00141212

To investigate how genetic variation is structured more finely within the basin populations, we analyzed groups of low and extra plated populations separately using STRUCTURE. In the analysis of the low plated grouping we found a model of $K = 3$ best fit the data. A plot of this level of K shows that Dougren Slough, Lynx Hollow Slough, and Jont Creek all separate into nearly three unique freshwater clusters. Finley Grey Creek Swamp, Green Island, and Science Factory represent a mixture of genotypes from these three clusters (Figure 3.2D). Interestingly, Dougren Slough, Lynx Hollow, and Jont Creek are all located near the heads of their individual drainages. Science Factory and

Green Island are located along the valley floor, low in the main stem of the Willamette, and downstream from Dougren Slough and Lynx Hollow. Finley Grey Creek Swamp is also located on the valley floor in a location that likely has high connectivity with the main stem via historical and present day flooding, (Benner and Sedell, 1997; O'Connor, 2001). The structuring of these populations is consistent with observations in other riverine fish systems of strong unidirectional gene flow in upstream populations presumably resulting from increasing river gradients with gain in elevation (Hernandez-Martich and Smith, 1990; Shaw et al., 1991; Castric et al., 2001).

Table 3.5. PCA table of analysis done with all of the Willamette Basin populations. Eigenvalues, percent variance explained, and p-values of PCs of genetic variation. The first PC accounts for the vast majority of the genetic variation and is the only axis that is statistically significantly.

	Eigenvalue	% Variance	Cumulative	*P*-value
1	28.545	56.37	56.37	0.03
2	6.58	12.993	69.363	0.92
3	4.266	8.425	77.788	0.96
4	4.12	8.136	85.924	0.59
5	2.396	4.731	90.655	0.97
6	1.545	3.051	93.707	1

The extra plated grouping demonstrated substructure. We found that a model of K = 3 best fit the data in a STRUCTURE analysis. Plots of this level of K show Leaburg and Milk Creek populations clustering independently from one another (Figure 3.2C). Walterville Slough and Riverbend have a mixture of some alleles shared with Leaburg and others most likely with one of the non-oceanic basin clusters. Buell-Miller Slough also appears to share alleles with Leaburg but these are admixed with alleles from a different non-oceanic basin cluster than Walterville Slough and Riverbend (Figure 3.2C).

Table 3.6. AMOVA table demonstrating partitioning of genetic variation between low and extra plated groupings.

Source of Variation	Nested in	% Variance	*F*-stat	P-value
Within Individual	-----	0.538	F_{IT}	-----
Among Individual	Population	0.051	F_{IS}	0.001
Among Population	Low vs Extra	0.28	F_{SC}	0.001
Among Low Extra	-----	0.13	F_{CT}	0.004

In STRUCTURE analysis, alternate models of *K* surrounding the best-fit model can be informative in understanding different levels of genetic structure in the populations under study, (Ostrander and Wayne, 2005). We investigated increasing levels of *K* in the full data set to understand contextually how the pattern of fine structure seen in the basin relates across subgroupings and how it relates to the broader sample of Oregon populations. We found nearly the same clusters at $K = 4$ in the full data set as we found in both analyses of subgroupings (Figure 3.1C). At this level of *K*, coastal populations (and Leaburg) are still clearly differentiated from the basin populations. Additionally, the partitioning of coastal and Willamette Basin populations is maintained in this analysis.

Interestingly, in the overall analysis at $K = 4$ (and in the subgrouping analysis) it appears that there are three freshwater clusters in the basin (Figure 3.1C, indicated in yellow, blue, and red). Also, it appears that these clusters may be spatially separated. Alleles from one cluster were only found in Northern basin populations, solely in Jont Creek and in mixed genotypes in Milk Creek, Buell-Miller, and Finley Creek Swamp (Figure 3.1C, 3.4, yellow cluster). The two other clusters are mainly found in the Southern basin. One is nearly solely present in Dougren Slough and is in mixed genotypes in the McKenzie, Finley Grey Creek Swamp, Green Island, and Science Factory (Figure 3.1C, 3.4, red cluster). The other is solely present in Lynx Hollow Slough

and is in mixed genotypes in Finley Grey Creek Swamp, Green Island, Science Factory, and Riverbend (Figure 3.1C, 3.4, blue cluster).

Unexpectedly, there is very strong genetic divergence among populations within the basin. Between the low plated populations F_{ST} values range from 0.07 (Science Factory and Green Island) to 0.43 (Dougren Slough and Lynx Hollow). This strong divergence is also observed between Jont Creek and Dougren Slough and Jont Creek and Lynx Hollow (~ 0.31) for both. Genetic divergence between the Leaburg population and the low plated populations is extreme with F_{ST} ranging from ~0.32 between Leaburg and Jont Creek to ~0.46 between Leaburg and Dougren Slough (Table 3.3).

Additionally, the basin populations closely group with Central Oregon populations concordant with what has been recently found (Catchen et al., 2013) and consistent with the hypothesis that the upper Deschutes Basin populations are the result of a recent introduction from the Willamette Basin (Figure 3.1C). From this analysis it appears that the Central Oregon populations originated from a subset of the alleles found in the Southern Willamette Basin.

Populations in the basin are partially isolated by distance

We found a signal of IBD in the Willamette Basin populations (Figure 3.5). Geographic distance and genetic divergence are moderately correlated (Mantel's r = 0.398, $P = 0.033$). However, to test whether the potential presence of transplanted oceanic fish within the basin could confound inference from the IBD result above, we performed a second analysis excluding the phenotypically oceanic populations Buell-Miller, Walterville, and Leaburg Lake. This resulted in a significant correlation between

geographical distance and genetic divergence (Mantel's r = 0.561, $P = 0.020$; data not shown).

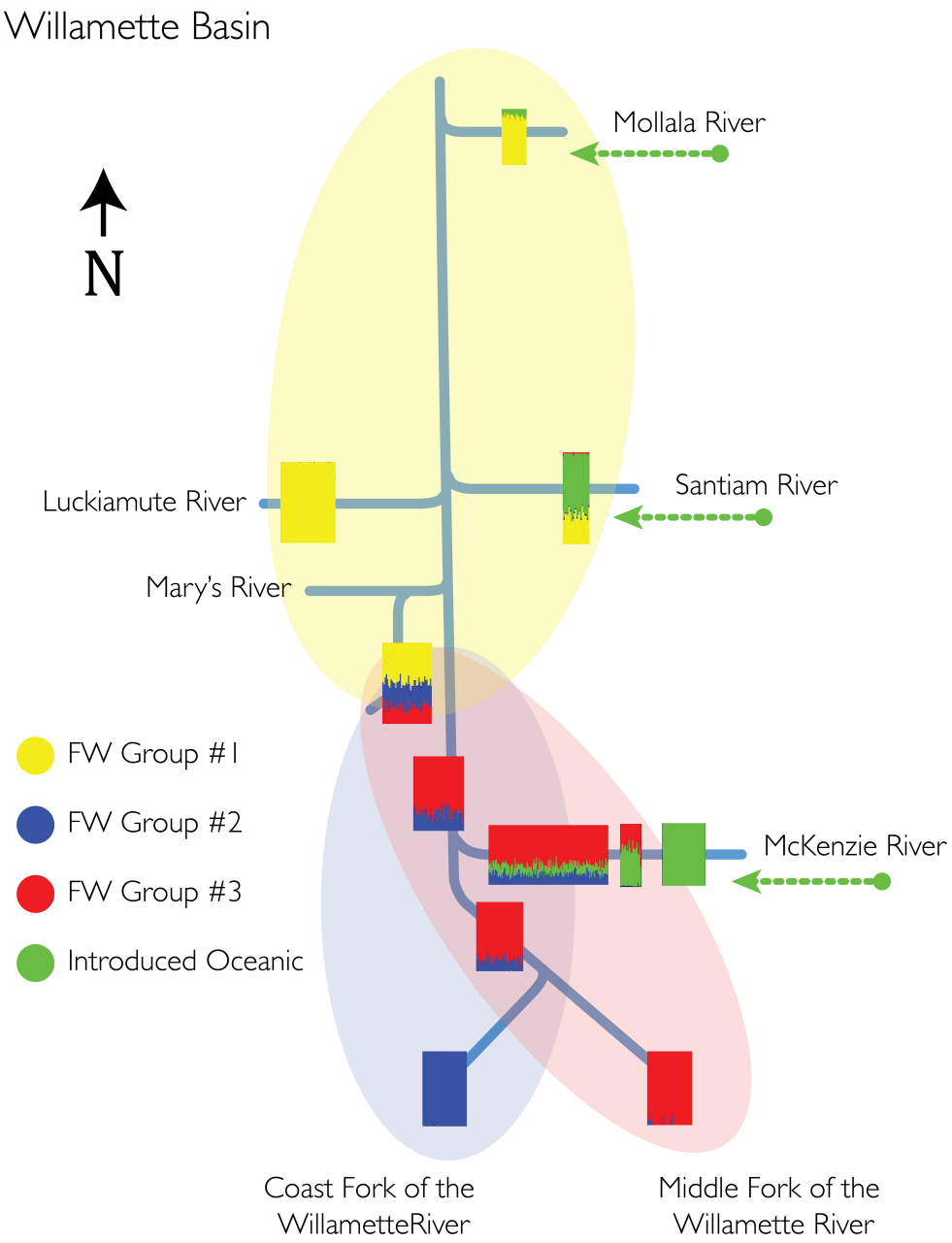

Figure 3.4. Subway plot of the major rivers in the Willamette Basin demonstrating the geographic distribution of genetic clusters found with STRUCTURE. Yellow genotypes are located in the north. Red and blue clusters are found in the south but appear to be almost fully partitioned between Coast and Middle Forks of the Willamette. Oceanic genotypes are represented in green and are differentially introgressed into freshwater genotypes.

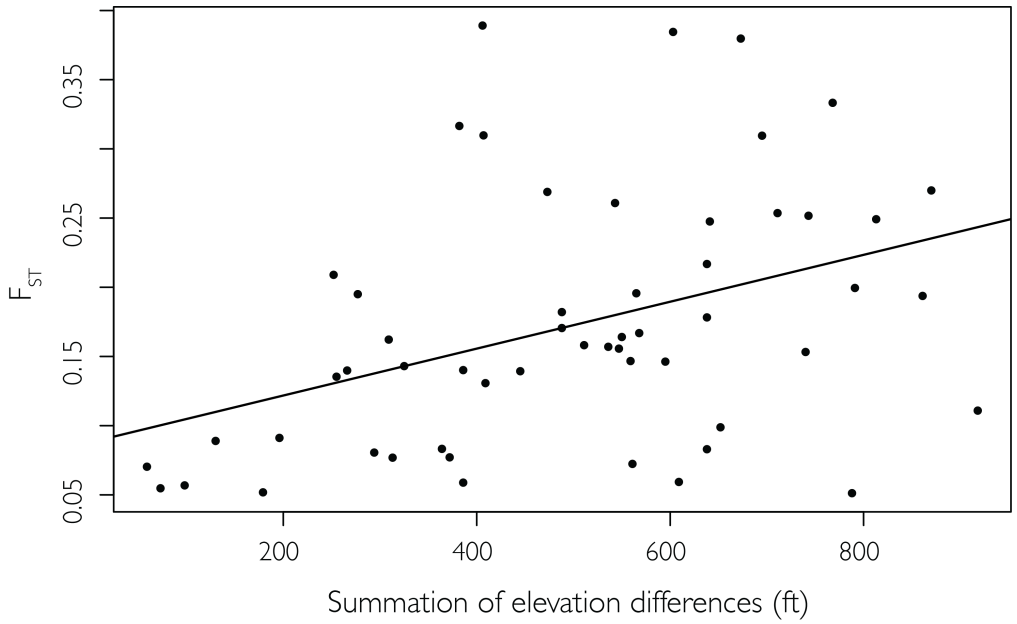

Figure 3.5. Analysis of IBD in the Willamette Basin populations using a Mantel's test. This test was done using pairwise genome-wide estimates of genetic divergence (F_{ST}) and Summation of elevation differences.

There is a severe reduction of nucleotide diversity in some low plated Willamette Basin populations

Nucleotide diversity in the basin populations varies greatly (Table 3.4). Values of π in low plated populations range from 0.00011 (Dougren Slough) to 0.00076 (Science Factory). These values are notably reduced from those that have been reported in Alaskan (Hohenlohe et al., 2010) and other Oregon (Catchen et al., 2013) populations. Lynx Hollow, Dougren Slough, and to some extent Jont Creek all have very low nucleotide diversity compared to the other populations within the basin, with genome wide nucleotide diversity of 0.0002, 0.0001, and 0.0004 respectively compared to the average of the other low plated populations 0.0007. High plated populations have noticeably greater nucleotide diversity and range from 0.0014 (Milk Creek) to 0.0026 (Buell-Miller Slough; Table 3.4).

DISCUSSION

The major partitioning of genetic variation is between coastal and inland populations despite phenotypic and life history groupings

The majority of the genetic variation in stickleback that inhabit the regions of Oregon that we studied is partitioned between coastal and inland populations (Figure 3.1). This partitioning exists despite phenotypic and life history groupings. Coastal low plated freshwater populations are more closely related to oceanic high plated populations than they are to inland freshwater low plated populations. These results suggest greater amounts of gene flow between the coastal fresh and oceanic populations than either has with the inland populations.

There is large amounts of divergence between coastal and Willamette Basin populations as demonstrated by F_{ST} analysis. This is concordant with what has been found in research with a subset of this data (Catchen et al., 2013) and what has been found in other stickleback systems, outside of Oregon, that are thought to predate the last ice age (Makinen et al., 2006). These findings support the hypothesis that Willamette Basin populations may be old and precede the last glacial maximum.

Geological evidence suggests that historically the Southern Willamette Basin drained directly west through the Siuslaw River and entered the ocean near present day Florence, Oregon. Approximately 1 million years ago this flow was captured by the main stem of the Willamette River and diverted northward up the Willamette Valley and out the Colombia River as it currently flows (Baldwin and Howell, 1949). Thus southern basin populations were once much closer to the sea (~200 km) than present day and the order of proximity to oceanic populations was inverted so that Cushman Slough and

South Jetty would have been the closest to populations in the southern basin. We found that populations in the southern basin are more closely related to southern oceanic populations than they are to northern oceanic populations in that genetic divergence between basin and oceanic populations decreases with increased distance (as the modern fish swims) along the coast. This pattern of divergence between the basin and oceanic populations could be a remnant of these geological events. A more fine scale survey of coastal and northern basin populations may help clarify this relationship.

The recent origin of phenotypically oceanic stickleback in the Willamette Basin

Genetic evidence supports the hypothesis of a recent introduction of oceanic fish in the Willamette Basin. The Leaburg Lake population clearly groups with the coastal populations in PC analysis and is solely made up of coastal genotypes in all groupings analyzed with STRUCTURE. This analysis also clearly demonstrates that Walterville and Buell-Miller closely group with coastal populations and are made up of mostly coastal genotypes (Figure 3.1A, B). Analysis of genetic divergence shows that Leaburg, Buell-Miller, and Walterville are more closely related to coastal populations than they are to low plated populations in the basin. To exemplify this point, Leaburg and Cushman Slough are ~ 770 km apart however they have very little genomic divergence (~0.026). Conversely, Walterville Slough is much more divergent from Leaburg (~ 0.11) which is ~ 21 km away, and the closest low plated population to Leaburg (Green Island), 63 km away, is very strongly diverged with F_{ST} ~ 0.37. Fish collection records at Leaburg also support the hypothesis of a recent introduction of oceanic stickleback. Intensive fish collections at Leaburg Dam in the 1980's found no stickleback (Zakel and Reed, 1984).

Phenotypic, genetic, and collecting evidence all strongly point to a very recent introduction, within the last 30 years, of oceanic threespine stickleback into Leaburg Lake. However, the life history of phenotypically oceanic and extra plated stickleback in other tributaries, Santiam and Mollala, is still unclear. The data indicate that oceanic genotypes have clearly introgressed into these populations, but the recent history of these alleles is unknown. Increased sampling of stickleback populations throughout these rivers and further investigations of collecting and stocking literature may help clarify this matter.

A hybrid zone exists in the McKenzie River

Previous research identified a phenotypic cline in the McKenzie suggestive of introgressive hybridization between low plated fish, occurring downstream and extra plated occurring upstream, with intermediate forms at some intervening sampling sites (Currey et al. (in prep)). Analysis of genetic variation clearly indicates that populations located between Green Island and Leaburg Lake are genetically intermediate (Figure 3.1A, 3.2A) with genotypes in these populations made up of both freshwater and recently introduced oceanic alleles (Figure 3.1B, 3.2B). There is also a nearly perfect relationship between river miles and the partitioning of genetic variation along this river (Figure 3.3A). Additionally, the partitioning of genetic variation nearly perfectly correlates with the partitioning of phenotypic variation (Figure 3.3B). The majority of the 1000 randomly chosen loci have presumably evolved neutrally, and we therefore expect that the relationships that we see between phenotypic and genetic variation are not the result of natural selection but are the product of other evolutionary processes. F_{IS} is relatively

elevated in the intermediate populations suggesting that non-random mating or some other form of cryptic genetic structure exists in these populations or that there has been recent hybridization. These phenotypic and genetic patterns seen in the McKenzie River are similar to what has been found in other known hybrid zones (Streisfeld and Kohn, 2005; Costedoat et al., 2007; Gay et al., 2008; Chavez et al., 2011). This evidence strongly supports the hypothesis of a hybrid zone in the McKenzie River.

Genetic variation is spatially structured in Willamette Basin threespine stickleback

The major axis of genetic structuring of stickleback in the Willamette Basin recapitulates what was found in the broader analysis including coastal populations. Populations are separated into one group that contains oceanic and freshwater alleles and another that contains only freshwater alleles, and these groups correspond to extra plated and low plated phenotypic groupings, respectively (Figure 3.2A, B). This is due to the recent introduction of oceanic stickleback in the eastern drainages of the basin and resulting introgression into subsets of the basin freshwater populations.

There is also substructure in the Willamette Basin populations of stickleback. It appears that there are at least three distinct genetic clusters that are spatially partitioned between northern and southern regions of the basin, and partitioned uniquely in upstream populations or mixed in populations located along the valley floor (Figure 3.4). Sampling efforts suggest that Jont Creek, Lynx Hollow, and Dougren Slough represent the upstream limits of stickleback inhabitance in each of their respective tributaries (Figure 3.4). Additionally, these upstream populations demonstrate very low levels of genetic diversity (Table 3.4).

These patterns of reduced genetic diversity, and unique partitioning of genetic variation, are consistent with what has been found in other riverine fish systems (Hernandez-Martich and Smith, 1990; Shaw et al., 1991; Castric et al., 2001). These patterns of genetic variation are also consistent with the hypothesis that populations high in river systems experience decreased probability of gene flow thought to be caused by physical restrictions of fish movement (water falls, river gradients, etc.) or due to founder effects caused by the decrease in the probability of the number of founders with increase in river miles (Castric et al., 2001). The relatively low levels of genetic diversity observed in these upstream populations support a hypothesis of reduced effective population size (N_e) via founder effects. However, it's unclear whether populations located high in individual drainages have differentially fixed standing genetic variation located in the mixed population lower in the basin or if these high populations represent old very diverged populations with subsequent migration and admixing in populations lower in the basin. Further sampling and investigation will be necessary to tease out these two scenarios.

We found a moderate signal of IBD that can partially explain the substructure that we observed in the basin populations. The moderate signal could indicate inadequacies of our geographical distance model, SOE, or that there are other life history and natural processes influencing these populations. The basin has a rich history of large flooding events caused by early snow followed by warming (Benner and Sedell, 1997) or massive flooding as seen during the Missoula Floods (O'Connor, 2001). Additionally, observational evidence suggests that stickleback within the basin may have migratory life

histories (however the extent to this migration is unknown). These differences in flow regime and the potentiality of migration suggests that connectivity between populations may be more complex than can be modeled using geographic distance measures.

Greater genetic divergence of threespine stickleback populations in Oregon than Alaska supports Oregon populations being older

There appears to be predictable patterns of genetic divergence in populations of very different ages (Figure 3.6). Very young freshwater populations located on Middleton Island show very little divergence from oceanic populations (Lescak et al. (in prep)). This divergence is greater in older south-central Alaskan freshwater populations compared to oceanic. Our data show that the divergence is even larger in very old Oregon freshwater populations compared to oceanic. However, distance from the ocean, life history differences, and geological differences could all influence these measures of divergence. For instance, freshwater populations within the Willamette Basin are much farther from the ocean than those found on Middleton Island or south-central Alaska. It is interesting to note that Leaburg Lake, which appears to be very recently introduced but spatially very distant from the ocean, has the same magnitude of divergence as very young Middleton Island populations suggesting that time since divergence is a major factor influencing genetic divergence.

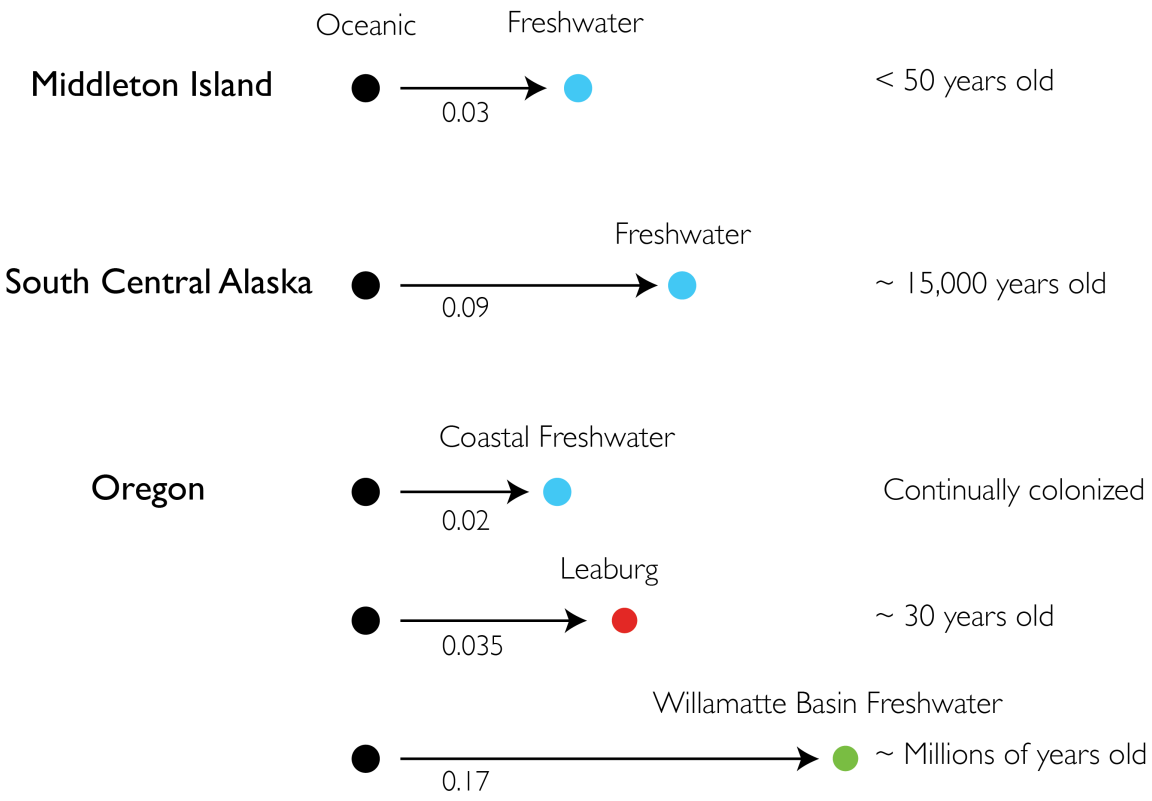

Figure 3.6. Summary of F_{ST} distances between varying aged stickleback systems. F_{ST} progressively grows, as the systems get older. Black dots represent oceanic populations. Blue dots represent non Willamette Basin freshwater populations. The red dot represents the recently introduced Leaburg population. The green dot represents the low plated freshwater populations in the Willamette Basin.

CONCLUSIONS

Here we have shown that the major structuring of genetic variation of Oregon stickleback is between coastal and inland populations. This suggests that there is greater gene flow between coastal freshwater and oceanic than either has with the inland populations. We have also shown that Willamette Basin populations are potentially much older than recently derived (<50 and 13,000 years old) freshwater populations of stickleback and that basin populations may have originated before the last ice age. Additionally, we may have identified a genomic signature in stickleback of an old geologic event that drastically changed the course of drainage in the Willamette Basin.

We have identified a recent introduction of oceanic stickleback in the eastern drainages of the basin and identified a hybrid zone that formed in the McKenzie River as the result of this introduction.

We have also shown that genetic variation in Willamette Basin stickleback populations is spatially structured both by geographic distance and river gradient but also primarily along a north-south geographic axis. This sets up an interesting scenario with the recent introduction of oceanic genotypes. These spatial patterns of genetic structuring and differential introgression of oceanic alleles implies that in each tributary along the eastern region of the Willamette Basin there could be parallel instances of recently introduced oceanic genomes hybridizing with similar but different combinations of potentially old basin freshwater genomes (Figure 3.4).

CHAPTER IV

CONCLUSION

SPATIAL DISTRIBUTION OF PHENOTYPIC AND GENETIC VARIATION OF OREGON THREESPINE STICKLEBACK

I have shown that most of the phenotypic and genetic variation is partitioned between oceanic and Willamette Basin stickleback populations. This partitioning of variation is similar to what has been observed in other stickleback systems. However I did find significant partitioning of phenotypic and genetic variation within and between the Willamette Basin populations. Despite the lack of regional segregation of phenotypic variation in the basin I did find localized segregation of phenotypic variation. Additionally, I found that genetic variation is geographically structured partially by IBD and partitioned into upstream regions of the major rivers but also along a north – south gradient.

There is an apparent disjunction between the partitioning of phenotypic and genetic variation within the Willamette Basin. While I found regional partitioning of genetic variation I did not find regional partitioning of phenotypic variation. The lack of finding partitioning of phenotypic variation between east and western regions does not rule out that there are other regional axes that partition phenotypic variation. However, the regions that I did test were chosen because they appear to demonstrate the largest differences in aquatic habitats. It is more likely that there are no regional differences in phenotypic variation and that the phenotypic divergence that does exist is partitioned at

the local level due to localized differences in aquatic habitats and predator-prey regimes. I did find evidence for this in differences in spine length, pelvic structure morphology, and gill raker counts between populations within the basin. These traits have been shown to be ecologically important and vary between freshwater forms in other systems. Regional divergence of genetic variation and the disjunction with the phenotypic variation is probably due to the nature of the markers that I used. The majority of the 1000 randomly chosen loci have presumably evolved neutrally so that the regional and geographical differences that I see are due to neutral processes. If indeed phenotypic differences that I found are due to localized adaptation this should be reflected in other regions of the genome responsible for these differences and should show quite different patterns of genetic divergence. In future, the full dataset of tens of thousands of RAD loci genotyped for this project may help uncover genomic regions associated with phenotypic differences.

I discovered the existence of phenotypically and genetically oceanic stickleback in the basin and the existence of a phenotypic and genetic cline in the McKenzie River. These fish are most likely the result of a recent oceanic introduction and subsequent dispersal of oceanic genotypes in the McKenzie River. This data and these results suggest that oceanic genotypes are diffusing into ancient freshwater genomes neutrally in a gradient manner along the McKenzie River, however it is unclear of the role that selection is having in the formation of this cline. Both processes of neutral diffusion and natural selection would act differently on different parts of the genome and future research is needed to tease this out.

GENOMIC SIGNATURES OF ANCIENT GEOLOGICAL EVENTS

The Umpqua and Siuslaw Rivers were once connected to the upper Willamette River but captured by westward flowing streams during the late Cenozoic and late Pleistocene respectively (Baldwin and Howell, 1949; Baldwin, 1959). There is biological evidence of these ancient connections in several fish species: chub, *Oregonichthys* (Markle et al., 1991), dace, *Rhinichtyhys* (Bisson and Reimers, 1977), and trout, *Oncorhynchus* (personal communication with ODFW). I found that upper Willamette populations of stickleback are genomically less divergent from stickleback along the Southern Oregon Coast near Florence than they are from Northern Oregon Coast, even though the northern coastal populations are geographically closer, as the fish swims. This may be the genomic imprint of this ancient connection between the Willamette and the Siuslaw. However, more work is needed to test this hypothesis and solidify these findings.

TEMPORAL DISTRIBUTION OF PHENOTYPIC AND GENETIC VARIATION OF OREGON THREESPINE STICKLEBACK

I have shown that a stereotypical phenotypic transition in freshwater stickleback populations evolving from marine ancestors is shared between young freshwater populations and those that likely originated before the last ice age. This finding suggests an ancient origin for standing genetic variation that appears to be continually reused during the oceanic to freshwater transition.

I have also shown that Willamette Basin populations are potentially much older than recently derived (<50 and 13,000 years old) freshwater populations of stickleback and that basin populations may have originated before the last ice age.

From this and previous work it appears that populations of threespine stickleback that inhabit Oregon demonstrate a wide range of evolutionary scenarios. First, there are very large differences in the timing of freshwater divergence in Oregon populations. Coastal freshwater populations demonstrate little divergence from oceanic populations and appear to have much gene flow with oceanic populations. Conversely, this work demonstrates that inland populations are much more diverged from oceanic populations and may be much older with very little gene flow with oceanic populations. These populations present interesting opportunities to investigate evolutionary processes over multiple time scales.

Additionally, populations within the Willamette Basin are phenotypically differentiated and may prove useful in studies of localized freshwater adaptation. Basin populations also demonstrate geographical partitioning of genetic variation and apparent isolation of populations in up river regions, both of which may reflect isolation by distance and/or the complicated dynamics between gene flow and river processes (e.g. river gradient).

Interestingly, it appears that there are at least two recent events (both within ~ the last 30 years) of introduction in Oregon stickleback populations. One of these introductions was of oceanic fish into the McKenzie River (and possibly into other eastern rivers of the basin) and the other was of freshwater Willamette Basin fish into the Deschutes Basin. Both of these introductions present unique scenarios that can be taken

advantage of, allowing us to compare and contrast phenotypic, genetic, and genomic changes after recent introduction in multiple systems and differing time scales. These introductions and potential hybrid zones (such as that in the McKenzie) can also be used to look for the genetic basis of ecologically important traits by taking advantage of the breakage of genomic linkage blocks that can hinder precision in association mapping studies.

Lastly, there is also an opportunity to investigate the role that ancient geological events have had in shaping the distribution of phenotypic and genetic variation. This could be done by looking at stickleback populations in and near the Siuslaw and Umpqua and comparing these fish to Willamette Basin and Columbia River fish. The nature of these relationships may shed light on the history of upper basin stickleback and of other aquatic species, and may help inform policies created to protect these species.

The stickleback populations that call Oregon home present us with an "evolutionary playground" that may help us better understand how evolutionary processes shape phenotypic, genetic, and genomic variation into the great variety of life.

REFERENCES CITED

Aldenhoven, JT, MA Miller, PS Corneli, and MD Shapiro. "Phylogeography of Ninespine Sticklebacks (Pungitius Pungitius) in North America: Glacial Refugia and the Origins of Adaptive Traits." *Mol Ecol* 19, no. 18 (2010): 4061–76.

Allendorf, FW, PA Hohenlohe, and G Luikart. "Genomics and the Future of Conservation Genetics." *Nature Reviews Genetics* 11, no. 10 (2010): 697–709.

Baird, NA, PD Etter, TS Atwood, MC Currey, AL Shiver, ZA Lewis, EU Selker, WA Cresko, and EA Johnson. "Rapid Snp Discovery and Genetic Mapping Using Sequenced Rad Markers." *PLoS One* 3, no. 10 (2008): e3376.

Baldwin, EM and PW Howell. "The Long Tom, a Former Tributary of the Siuslaw River." *Northwest Science* 23 (1949): 112–24.

Baldwin, EM. *Geology of Oregon*. University of Oregon Cooperative Book Store Eugene, OR, 1959.

Barrett, RD and D Schluter. "Adaptation From Standing Genetic Variation." *Trends Ecol Evol* 23, no. 1 (2008): 38–44.

Bates, Henry Walter. "Notes on Coleopterous Insects Frequenting Damp Places." *Zoologist* 1 (1843): 114–15.

Bell, MA. "Lateral Plate Evolution in the Threespine Stickleback: Getting Nowhere Fast." *Genetica* 112, no. 1 (2001): 445–61.

Bell, MA, WE Aguirre, and NJ Buck. "Twelve Years of Contemporary Armor Evolution in a Threespine Stickleback Population." *Evolution* 58, no. 4 (2004): 814–24.

Bell, MA and G Ortí. "Pelvic Reduction in Threespine Stickleback From Cook Inlet Lakes: Geographical Distribution and Intrapopulation Variation." *Copeia* (1994): 314–25.

Bell, MA and SA Foster. *The Evolutionary Biology of the Threespine Stickleback*. Oxford; New York: Oxford University Press, 1994.

Benner, PA and JR Sedell. "Upper Willamette River Landscape: A Historic Perspective." *River quality: dynamics and restoration. CRC Press, Lewis Publishers, New York* (1997): 23–47.

Bisson, PA and PE Reimers. "Geographic Variation Among Pacific Northwest Populations of Longnose Dace, Rhinichthys-Cataractae." *Copeia* (1977): 518–22.

Boothl, DB, KG Troostl, JJ Clague, and RB Waitt. “The Cordilleran Ice Sheet.” *The Quaternary Period in the United States* 1 (2003): 17.

Bryc, K, A Auton, MR Nelson, JR Oksenberg, SL Hauser, S Williams, A Froment, JM Bodo, C Wambebe, SA Tishkoff, and CD Bustamante. “Genome-Wide Patterns of Population Structure and Admixture in West Africans and African Americans.” *Proc Natl Acad Sci U S A* 107, no. 2 (2010): 786–91.

Bustamante, CD, A Fledel-Alon, S Williamson, R Nielsen, MT Hubisz, S Glanowski, DM Tanenbaum, TJ White, JJ Sninsky, RD Hernandez, D Civello, MD Adams, M Cargill, and AG Clark. “Natural Selection on Protein-Coding Genes in the Human Genome.” *Nature* 437, no. 7062 (2005): 1153–57.

Butlin, RK, M Saura, G Charrier, B Jackson, C Andre, A Caballero, JA Coyne, J Galindo, JW Grahame, J Hollander, P Kemppainen, M Martinez-Fernandez, M Panova, H Quesada, K Johannesson, and E Rolan-Alvarez. “Parallel Evolution of Local Adaptation and Reproductive Isolation in the Face of Gene Flow.” *Evolution* 68, no. 4 (2014): 935–49.

Byrne, S, A Czaban, B Studer, F Panitz, C Bendixen, and T Asp. “Genome Wide Allele Frequency Fingerprints (Gwaffs) of Populations Via Genotyping By Sequencing.” *PLoS One* 8, no. 3 (2013): e57438.

Caicedo, AL, SH Williamson, RD Hernandez, A Boyko, A Fledel-Alon, TL York, NR Polato, KM Olsen, R Nielsen, SR McCouch, CD Bustamante, and MD Purugganan. “Genome-Wide Patterns of Nucleotide Polymorphism in Domesticated Rice.” *PLoS Genet* 3, no. 9 (2007): 1745–56.

Castric, V, F Bonney, and L Bernatchez. “Landscape Structure and Hierarchical Genetic Diversity in the Brook Charr, Salvelinus Fontinalis.” *Evolution* 55, no. 5 (2001): 1016–28.

Catchen, J, S Bassham, T Wilson, M Currey, C O’Brien, Q Yeates, and WA Cresko. “The Population Structure and Recent Colonization History of Oregon Threespine Stickleback Determined Using Restriction-Site Associated DNA-Sequencing.” *Mol Ecol* 22, no. 11 (2013): 2864–83.

Catchen, JM, A Amores, P Hohenlohe, W Cresko, and JH Postlethwait. “Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences.” *G3 (Bethesda)* 1, no. 3 (2011): 171–82.

Chavez, AS, CJ Saltzberg, and GJ Kenagy. “Genetic and Phenotypic Variation Across a Hybrid Zone Between Ecologically Divergent Tree Squirrels (Tamiasciurus).” *Mol Ecol* 20, no. 16 (2011): 3350–66.

Colosimo, PF, KE Hosemann, S Balabhadra, G Jr Villarreal, M Dickson, J Grimwood, J Schmutz, RM Myers, D Schluter, and DM Kingsley. "Widespread Parallel Evolution in Sticklebacks By Repeated Fixation of Ectodysplasin Alleles." *Science* 307, no. 5717 (2005): 1928–33.

Costedoat, C, N Pech, R Chappaz, and A Gilles. "Novelties in Hybrid Zones: Crossroads Between Population Genomic and Ecological Approaches." *PLoS One* 2, no. 4 (2007): e357.

Cresko, WA, A Amores, C Wilson, J Murphy, M Currey, P Phillips, MA Bell, CB Kimmel, and JH Postlethwait. "Parallel Genetic Basis for Repeated Evolution of Armor Loss in Alaskan Threespine Stickleback Populations." *Proceedings of the National Academy of Sciences of the United States of America* 101, no. 16 (2004): 6050–55.

Darwin, CR. *On the Origin of Species By Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. 1st ed. John Murray, London 1859.

Davey, JW, PA Hohenlohe, PD Etter, JQ Boone, JM Catchen, and ML Blaxter. "Genome-Wide Genetic Marker Discovery and Genotyping Using Next-Generation Sequencing." *Nature Reviews Genetics* 12, no. 7 (2011): 499–510.

Day, T, J Pritchard, and D Schluter. "A Comparison of Two Sticklebacks." *Evolution* (1994): 1723–34.

Deagle, BE, FC Jones, YF Chan, DM Absher, DM Kingsley, and TE Reimchen. "Population Genomics of Parallel Phenotypic Evolution in Stickleback Across Stream-Lake Ecological Transitions." *Proc Biol Sci* 279, no. 1732 (2012): 1277–86.

Dobzhansky, T. "A Review of Some Fundamental Concepts and Problems of Population Genetics." Cold Spring Harbor Symposia on Quantitative Biology 20 (1955): 1–15.

Dobzhansky, T. *Genetics and the Origin of Species*. Columbia University Press, 1937.

Earl, DA. "Structure Harvester: A Website and Program for Visualizing Structure Output and Implementing the Evanno Method." *Conservation genetics resources* 4, no. 2 (2012): 359–61.

Emerson, KJ, CR Merz, JM Catchen, PA Hohenlohe, WA Cresko, WE Bradshaw, and CM Holzapfel. "Resolving Postglacial Phylogeography Using High-Throughput Sequencing." *Proc Natl Acad Sci U S A* 107, no. 37 (2010): 16196–200.

Endler, JA. *Geographic Variation, Speciation, and Clines*. Princeton University Press, 1977.

Endler, JA. “On the Measurement and Classification of Colour in Studies of Animal Colour Patterns.” *Biological Journal of the Linnean Society* 41, no. 4 (1990): 315–52.

Endler, JA. *Natural Selection in the Wild*. Princeton, N.J.: Princeton University Press, 1986.

Etter, PD, S Bassham, PA Hohenlohe, EA Johnson, and WA Cresko. “SNP Discovery and Genotyping for Evolutionary Genetics Using Rad Sequencing,” In *Molecular Methods for Evolutionary Genetics*, 157–78. Springer, 2011.

Evanno, G, S Regnaut, and J Goudet. “Detecting the Number of Clusters of Individuals Using the Software Structure: A Simulation Study.” *Mol Ecol* 14, no. 8 (2005): 2611–20.

Excoffier, L, PE Smouse, and JM Quattro. “Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data.” *Genetics* 131, no. 2 (1992): 479–91.

Falush, D, M Stephens, and JK Pritchard. “Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies.” *Genetics* 164, no. 4 (2003): 1567–87.

Fisher, RA. *The Genetical Theory of Natural Selection*. Рипол Классик, 1958.

Gay, L, PA Crochet, DA Bell, and T Lenormand. “Comparing Clines on Molecular and Phenotypic Traits in Hybrid Zones: A Window on Tension Zone Models.” *Evolution* 62, no. 11 (2008): 2789–806.

Gelmond, O, FA von Hippel, and MS Christy. “Rapid Ecological Speciation in Three-Spined Stickleback Gasterosteus Aculeatus From Middleton Island, Alaska: The Roles of Selection and Geographic Isolation.” *J Fish Biol* 75, no. 8 (2009): 2037–51.

Glenn, TC. “Field Guide to Next-Generation DNA Sequencers.” *Mol Ecol Resour* 11, no. 5 (2011): 759–69.

Grant, BR and PR Grant. “Fission and Fusion of Darwin’s Finches Populations.” *Philos Trans R Soc Lond B Biol Sci* 363, no. 1505 (2008): 2821–29.

Hagen, DW and LG Gilbertson. “Geographic Variation and Environmental Selection in Gasterosteus Aculeatus L. In the Pacific Northwest, America.” *Evolution* (1972): 32–51.

Haldane, JBS. “The Theory of a Cline.” *Journal of genetics* 48, no. 3 (1948): 277–84.

Haldane, JBS. “Population Genetics.” *New Biology* 18 (1955): 34–51.

Hartl, DL and AG Clark. *Principles of Population Genetics*. Sinauer associates Sunderland, 1997.

Hendry, AP, K Hudson, JA Walker, K Rasanen, and LJ Chapman. “Genetic Divergence in Morphology-Performance Mapping Between Misty Lake and Inlet Stickleback.” *J Evol Biol* 24, no. 1 (2011): 23–35.

Hernandez-Martich, JD and MH Smith. “Patterns of Genetic Variation in Eastern Mosquitofish (Gambusia Holbrooki Girard) From the Piedmont and Coastal Plain of Three Drainages.” *Copeia* (1990): 619–30.

Hoekstra, HE, KE Drumm, and MW Nachman. “Ecological Genetics of Adaptive Color Polymorphism in Pocket Mice: Geographic Variation in Selected and Neutral Genes.” *Evolution* 58, no. 6 (2004): 1329–41.

Hohenlohe, PA, S Bassham, PD Etter, N Stiffler, EA Johnson, and WA Cresko. “Population Genomics of Parallel Adaptation in Threespine Stickleback Using Sequenced Rad Tags.” *PLoS Genet* 6, no. 2 (2010): e1000862.

Hohenlohe, PA, J Catchen, and WA Cresko. “Population Genomic Analysis of Model and Nonmodel Organisms Using Sequenced Rad Tags,” In *Data Production and Analysis in Population Genomics*, 235–60. Springer, 2012.

Hoogland, R, D Morris, and N Tinbergen. “The Spines of Sticklebacks (Gasterosteus and Pygosteus) as Means of Defence Against Predators (Perca and Esox).” *Behaviour* (1956): 205–36.

Hughes, RM and JR Gammon. “Longitudinal Changes in Fish Assemblages and Water Quality in the Willamette River, Oregon.” *Transactions of the American Fisheries Society* 116, no. 2 (1987): 196–209.

Jefferson, A, GE Grant, and SL Lewis. “A River Runs Underneath it: Geological Control of Spring and Channel Systems and Management Implications, Cascade Range, Oregon.” *US Forest Service Pacific Northwest Research Station General Technical Report PNW-GTR (689, Part 2)* (2007): 391–400.

Jones, AW, EP Palkovacs, and DM Post. “Recent Parallel Divergence in Body Shape and Diet Source of Alewife Life History Forms.” *Evol Ecol* 27, no. 6 (2013): 1175–87.

Jones, FC, MG Grabherr, YF Chan, P Russell, E Mauceli, J Johnson, R Swofford, M Pirun, MC Zody, S White, E Birney, S Searle, J Schmutz, J Grimwood, MC Dickson, RM Myers, CT Miller, BR Summers, AK Knecht, SD Brady, H Zhang, AA Pollen, T Howes, C Amemiya, J Baldwin, T Bloom, DB Jaffe, R Nicol, J Wilkinson, ES Lander, F Di Palma, K Lindblad-Toh, and DM Kingsley. "The Genomic Basis of Adaptive Evolution in Threespine Sticklebacks." *Nature* 484, no. 7392 (2012): 55–61.

Kalbe, M and J Kurtz. "Local Differences in Immunocompetence Reflect Resistance of Sticklebacks Against the Eye Fluke Diplostomum Pseudospathaceum." *Parasitology* 132, no. Pt 1 (2006): 105–16.

Kitano, J, DI Bolnick, DA Beauchamp, MM Mazur, S Mori, T Nakano, and CL Peichel. "Reverse Evolution of Armor Plates in the Threespine Stickleback." *Curr Biol* 18, no. 10 (2008): 769–74.

Lavin, PA and JD McPhail. "Adaptive Divergence of Trophic Phenotype Among Freshwater Populations of the Threespine Stickleback (Gasterosteus Aculeatus)." *Canadian Journal of Fisheries and Aquatic Sciences* 43, no. 12 (1986): 2455–63.

Lescak, EA, FA von Hippel, RR Bernhardt, and MA Bell. "Pelvic Girdle Reduction and Asymmetry in Threespine Stickleback From Wallace Lake, Alaska." *Evol. Ecol. Res.* 15, no. 2 (2013): 155–70.

Lucek, K, A Sivasundar, BK Kristjansson, S Skulason, and O Seehausen. "Quick Divergence But Slow Convergence During Ecotype Formation in Lake and Stream Stickleback Pairs of Variable Age." *J Evol Biol* (2014):

Makinen, HS, JM Cano, and J Merila. "Genetic Relationships Among Marine and Freshwater Populations of the European Three-Spined Stickleback (Gasterosteus Aculeatus) Revealed By Microsatellites." *Mol Ecol* 15, no. 6 (2006): 1519–34.

Markle, DF, TN Pearsons, and DT Bills. "Natural History of Oregonichthys (Pisces: Cyprinidae), With a Description of a New Species From the Umpqua River of Oregon." *Copeia* (1991): 277–93.

Mcphail, JD. "Ecology and Evolution of Sympatric Sticklebacks (Gasterosteus) - Morphological and Genetic-Evidence for a Species Pair in Enos Lake, British-Columbia." *Can. J. Zool.-Rev. Can. Zool.* 62, no. 7 (1984): 1402–8.

Meirmans, PG and PH Van Tienderen. "Genotype and Genodive: Two Programs for the Analysis of Genetic Diversity of Asexual Organisms." *Mol Ecol Notes* 4, no. 4 (2004): 792–94.

Millet, A, BK Kristjansson, A Einarsson, and K Rasanen. “Spatial Phenotypic and Genetic Structure of Threespine Stickleback (Gasterosteus Aculeatus) in a Heterogeneous Natural System, Lake Myvatn, Iceland.” *Ecol Evol* 3, no. 10 (2013): 3219–32.

Münzing, J. “The Evolution of Variation and Distributional Patterns in European Populations of the Three-Spined Stickleback, Gasterosteus Aculeatus.” *Evolution* (1963): 320–32.

Nei, M. *Molecular Population Genetics and Evolution.* North-Holland Publishing Company., 1975.

Nosil, P and JL Feder. “Genomic Divergence During Speciation: Causes and Consequences.” *Philos Trans R Soc Lond B Biol Sci* 367, no. 1587 (2012): 332–42.

O’Connor, JE. *Origin, Extent, and Thickness of Quaternary Geologic Units in the Willamette Valley, Oregon*. US Geological Survey, 2001.

O’Quin, KE, CM Hofmann, HA Hofmann, and KL Carleton. “Parallel Evolution of Opsin Gene Expression in African Cichlid Fishes.” *Mol Biol Evol* 27, no. 12 (2010): 2839–54.

Olafsdottir, GA, SS Snorrason, and MG Ritchie. “Postglacial Intra-Lacustrine Divergence of Icelandic Threespine Stickleback Morphs in Three Neovolcanic Lakes.” *J Evol Biol* 20, no. 5 (2007): 1870–81.

Ostrander, EA and RK Wayne. “The Canine Genome.” *Genome Res* 15, no. 12 (2005): 1706–16.

Andrus, C, W Pearcy. “Temperature Monitoring and Modeling of the Mary’s River Watershed.” (1999):

Pritchard, JK, M Stephens, and P Donnelly. “Inference of Population Structure Using Multilocus Genotype Data.” *Genetics* 155, no. 2 (2000): 945–59.

Protas, ME, C Hersey, D Kochanek, Y Zhou, H Wilkens, WR Jeffery, LI Zon, R Borowsky, and CJ Tabin. “Genetic Analysis of Cavefish Reveals Molecular Convergence in the Evolution of Albinism.” *Nat Genet* 38, no. 1 (2006): 107–11.

R Development Core Team. “R: A Language and Environment for Statistical Computing.” R Foundation for Statistical Computing, Vienna, Austria (2013).

Reimchen, TE. “Structural Relationships Between Spines and Lateral Plates in Threespine Stickleback (Gasterosteus Aculeatus).” *Evolution* 37 (1983): 931–46.

Reimchen, TE and P Nosil. “Replicated Ecological Landscapes and the Evolution of Morphological Diversity Among Populations From an Archipelago on the West Coast of Canada.” *Can. J. Zool.* 84, no. 5 (2006): 643–54.

Reimchen, TE, EM Stinson, and JS Nelson. “Multivariate Differentiation of Parapatric and Allopatric Populations of Threespine Stickleback in the Sangan River Watershed, Queen-Charlotte-islands.” *Can. J. Zool.-Rev. Can. Zool.* 63, no. 12 (1985): 2944–51.

Reimchen, TE, C Bergstrom, and P Nosil. “Natural Selection and the Adaptive Radiation of Haida Gwaii Stickleback.” *Evolutionary Ecology Research* 15, no. 3 (2013): 241–69.

Reist, JD. “An Empirical Evaluation of Several Univariate Methods That Adjust for Size Variation in Morphometric Data.” *Canadian Journal of Zoology* 63, no. 6 (1985): 1429–39.

Reznick, DN, M Mateos, and MS Springer. “Independent Origins and Rapid Evolution of the Placenta in the Fish Genus Poeciliopsis.” *Science* 298, no. 5595 (2002): 1018–20.

Rosenblum, EB, H Rompler, T Schoneberg, and HE Hoekstra. “Molecular and Functional Basis of Phenotypic Convergence in White Lizards At White Sands.” *Proc Natl Acad Sci U S A* 107, no. 5 (2010): 2113–17.

Rundle, H. D. “Natural Selection and Parallel Speciation in Sympatric Sticklebacks.” *Science* 287, no. 5451 (2000): 306–8.

Rutter, C. *Notes on Fresh Water Fishes on the Pacific Slope of North America*. Leland Stanford Jr. University, 1896.

Schluter, D. “Adaptive Radiation in Sticklebacks - Size, Shape, and Habitat Use Efficiency.” *Ecology* 74, no. 3 (1993): 699–709.

Schluter, D. *The Ecology of Adaptive Radiation*. Oxford University Press, 2000.

Schluter, D and GL Conte. “Genetics and Ecological Speciation.” *Proceedings of the National Academy of Sciences* 106, no. Supplement 1 (2009): 9955–62.

Schluter, D and A Rambaut. “Ecological Speciation in Postglacial Fishes [and Discussion].” *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 351, no. 1341 (1996): 807–14.

Schneider, CA, WS Rasband, KW Eliceiri, J Schindelin, I Arganda-Carreras, E Frise, V Kaynig, M Longair, T Pietzsch, and S Preibisch. “Nih Image to Imagej: 25 Years of Image Analysis.” *Nature methods* 9, no. 7 (2012):

Seehausen, O. “African Cichlid Fish: A Model System in Adaptive Radiation Research.” *Proc Biol Sci* 273, no. 1597 (2006): 1987–98.

Shapiro, MD, MA Bell, and DM Kingsley. “Parallel Genetic Origins of Pelvic Reduction in Vertebrates.” *Proceedings of the National Academy of Sciences* 103, no. 37 (2006): 13753–58.

Shapiro, MD, ME Marks, CL Peichel, BK Blackman, KS Nereng, B Jónsson, D Schluter, and DM Kingsley. “Genetic and Developmental Basis of Evolutionary Pelvic Reduction in Threespine Sticklebacks.” *Nature* 428, no. 6984 (2004): 717–23.

Shaw, PW, GR Carvalho, AE Magurran, and BH Seghers. “Population Differentiation in Trinidadian Guppies (Poecilia Reticulata): Patterns and Problems.” *Journal of Fish Biology* 39, no. sA (1991): 203–9.

Shendure, J and H Ji. “Next-Generation DNA Sequencing.” *Nat Biotechnol* 26, no. 10 (2008): 1135–45.

Sokal, RR, GM Jacquez, and MC Wooten. “Spatial Autocorrelation Analysis of Migration and Selection.” *Genetics* 121, no. 4 (1989): 845–55.

Stacklies, W, H Redestig, M Scholz, D Walther, and J Selbig. “Pcamethods - a Bioconductor Package Providing Pca Methods for Incomplete Data.” *Bioinformatics* 23, no. 9 (2007): 1164–67.

Steiner, CC, JN Weber, and HE Hoekstra. “Adaptive Variation in Beach Mice Produced By Two Interacting Pigmentation Genes.” *PLoS biology* 5, no. 9 (2007): e219.

Storz, JF. “Using Genome Scans of DNA Polymorphism to Infer Adaptive Population Divergence.” *Mol. Ecol.* 14, no. 3 (2005): 671–88.

Streisfeld, MA and JR Kohn. “Contrasting Patterns of Floral and Molecular Variation Across a Cline in Mimulus Aurantiacus.” *Evolution* 59, no. 12 (2005): 2548–59.

Tague, C, M Farrell, G Grant, S Lewis, and S Rey. “Hydrogeologic Controls on Summer Stream Temperatures in the Mckenzie River Basin, Oregon.” *Hydrol. Process.* 21, no. 24 (2007): 3288–300.

Tague, C and GE Grant. “A Geological Framework for Interpreting the Low-Flow Regimes of Cascade Streams, Willamette River Basin, Oregon.” *Water Resour. Res.* 40, no. 4 (2004).

Vonholdt, BM, JP Pollinger, KE Lohmueller, E Han, HG Parker, P Quignon, JD Degenhardt, AR Boyko, DA Earl, A Auton, A Reynolds, K Bryc, A Brisbin, JC Knowles, DS Mosher, TC Spady, A Elkahloun, E Geffen, M Pilot, W Jedrzejewski, C Greco, E Randi, D Bannasch, A Wilton, J Shearman, M Musiani, M Cargill, PG Jones, Z Qian, W Huang, ZL Ding, YP Zhang, CD Bustamante, EA Ostrander, J Novembre, and RK Wayne. "Genome-Wide Snp and Haplotype Analyses Reveal a Rich History Underlying Dog Domestication." *Nature* 464, no. 7290 (2010): 898–902.

Waite, IR and KD Carpenter. "Associations Among Fish Assemblage Structure and Environmental Variables in Willamette Basin Streams, Oregon." *Transactions of the American Fisheries Society* 129, no. 3 (2000): 754–70.

Wallace, RA. *Contributions to the Theory of Natural Selection*. Macmillan, 1871.

Wallick, JR, GE Grant, ST Lancaster, JP Bolte, and RP Denlinger. "Patterns and Controls on Historical Channel Change in the Willamette River, Oregon, USA," 491–516. 2007.

Webster, MM, N Atton, PJ Hart, and AJ Ward. "Habitat-Specific Morphological Variation Among Threespine Sticklebacks (Gasterosteus Aculeatus) Within a Drainage Basin." *PLoS One* 6, no. 6 (2011): e21060.

Wilson, EO. *The insect societies.* Harvard University Press, 1971.

Wright, S. "Statistical Genetics and Evolution." *Bulletin of the American Mathematical Society* 48, no. 4 (1942): 223–46.

Wright, S. "An Analysis of Local Variability of Flower Color in Linanthus Parryae." *Genetics* 28, no. 2 (1943a): 139.

Wright, S. "Isolation By Distance." *Genetics* 28, no. 2 (1943b): 114.

Wright, S. *Evolution and the Genetics of Populations: A Treatise in Four Volumes: Vol. 4: Variability Within and Among Natural Populations*. University of Chicago Press, 1978.

Wu, TD and S Nacu. "Fast and Snp-Tolerant Detection of Complex Variants and Splicing in Short Reads." *Bioinformatics* 26, no. 7 (2010): 873–81.

Zakel, JC and DW Reed. "Downstream Migration of Fish At Leaburg Dam, Mckenzie River, Oregon, 1980 to 1983. Oregon Department of Fish and Wildlife Information Reports 84-13. Fish Division." *Research and Development Section, Corvallis* (1984):