UNIVERSITY OF CALIFORNIA
SANTA CRUZ

GENETIC INVESTIGATION OF THE PACIFIC TROUT COMPLEX: FROM PEDIGREES TO PHYLOGENIES

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

OCEAN SCIENCES

by

Alicia Abadía-Cardoso

March 2014

The Dissertation of Alicia Abadía-Cardoso
is approved:

John Carlos Garza, Ph.D., Chair

Jonathan Zehr, Ph.D.

Grant Pogson, Ph.D.

Francisco García de León, Ph.D.

Dean Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

GENETIC INVESTIGATION OF THE PACIFIC TROUT COMPLEX: FROM PEDIGREES TO PHYLOGENIES

by

Alicia Abadía-Cardoso

Perhaps one of the world's most important groups of fish are the species within the Pacific trout complex, due to their extensive harvest in fisheries and use in aquaculture. This dissertation consists of an in-depth evaluation of this group of trout at different scales, from the assessment of biological traits throughout the reconstruction of pedigrees in two trout populations to a phylogeographic examination of multiple largely undescribed native species. Here, a combination of novel molecular techniques allowed me to address critical ecological questions for the appropriate management and conservation of this group.

In chapter one, I describe how I discovered, characterized and developed a large number of single nucleotide polymorphisms (SNPs) for *O. mykiss* that allow study of ecological interactions, phylogeography, and conservation status. These molecular tools have great power for traditional population genetic analysis, and for individual identification and pedigree reconstruction. The last allows the tracking of families, and an unprecedented level of evaluation of natural and hatchery populations.

Chapter two expands the scope of these molecular tools to provide a powerful means of understanding of biological traits for steelhead hatchery programs in the

Russian River, California. Reconstruction of cohort age distributions revealed a strong component of fish that spawn at age two, in contrast to program goals and distinct from naturally spawning steelhead in the region. Correlations between family members in the day of spawning revealed a strongly heritable component to this life history trait and demonstrated the potential for selection to alter life history traits. These results demonstrate the promise of SNP-based pedigree reconstruction for providing biological inference in high-fecundity organisms.

Chapter three describes how the powerful molecular tools developed were applied to population genetics of trout inhabiting northwestern Mexico, to elucidate biodiversity, evaluate hypotheses regarding evolutionary history, and measure introgression from exotic hatchery rainbow trout. Here, I confirmed the vast genetic diversity present in northwestern Mexican trout and provided evidence that trout inhabiting the Sierra Madre Occidental correspond to independent lineages separated from *O. mykiss*. Introgression from non-native trout was detected, but the genetic integrity of native trout is still maintained in many watersheds. All the information presented in this document will help to guide effective conservation strategies for this globally important group of fishes.

A mis queridos abuelos

José, Trini,

Fernando y Martha.

Acknowledgments

First of all, I want to express my sincere gratitude to my PhD advisor Carlos Garza. None of this would have been possible without his immense support. The supervision and guidance he provided me as his student was essential in my path to become a researcher. His insightful and broad knowledge enlightened me enormously. He gave me the right pushes and was patient in the right moments, to just make me a better scientist.

I am especially thankful to Devon Pearse for his invaluable advise. I am grateful for all those times I spent discussing science and for always taking the time to listen to my questions and concerns about my work, or simply about life.

Over the years, I needed the assistance from many people in the various tasks I have undertaken. I want to thank Libby Gilbert-Horvath for all her help in the lab, making sure that everything worked properly and that I always had the necessary to run my samples, and Eric Anderson for always taking the time to assist me on any math/computer issues I encounter and for his invaluable collaboration on chapter 2.

I would also want to thank my past and present office and lab mates, with whom I spent endless hours sharing four walls. Most importantly among them was Anthony Clemento, for all the constructive discussions I had to make our research better.

This study would not have been possible without the sampling and associated efforts of multiple people to whom I am grateful. Thanks to Francisco García de León, Richard Mayden, and all the people from the binational group "Truchas Mexicanas" for providing the Mexican trout samples. Thanks to past and present members of the Molecular Ecology and Genetic Analysis Team: V. Apkenas, M. Arciniega, A. Clemento, C. Columbus, E. Gilbert-Horvath, H. Starks, D. Pearse A. Martínez, and E. Martínez for assistance with laboratory analyses and comments. I also thank the staff at WSH and CVFF, including P. LaCivita, E. McKenna, R. Taylor, B. White and B. Wilson. Additional thanks to A. Aguilar, N. Campbell, G. Charrier, C. Donohoe, G. Edwards, H. Fish, S. Hayes, A. Matala, L. Seeb and P. Moran for help with protocol development, supplying samples or insightful discussions.

I also need to thank my committee members Jonathan Zher, Grant Pogson, and Francisco García de León for their valuable comments to my proposal and especially in the final stages of this dissertation.

In a personal note, I want to thank my mom, my dad, and my brother for their endless encouragement from the distance.

Many thanks to my terrific friends that made this endeavor more bearable and fun. The Mexicans: Martha, Horacio, Stella, Sergio, and Emilio...and Luis. The dinner cluppers: Sally, Jimmy, Alice, Cyril, and lately Anne Marie and Bill. My (putative)

cohort: Melinda, Justin C., Jimmy, Cyril and Justin Y. The soccermates: all of them...

Last but not least, I want to express my deep gratitude to my partner-for-life Rodrigo. Once I heard that if you want to really meet someone you have to travel with that person. A few years ago, Rodrigo and I decided to take this arduous but gratifying trip that is the PhD and I have learn together not only about science but more importantly, about shared life and love.

Funding was provided by the CONACyT-UC Mexus Program, the National Marine Fisheries Service, Southwest Fisheries Science Center, and the US Army Corps of Engineers. The text of this dissertation includes reprints of the following previously published material:

Abadía-Cardoso A, Clemento AJ, Garza JC (2011) Discovery and characterization of single nucleotide polymorphisms in steelhead/rainbow trout, *Oncorhynchus mykiss*. *Molecular Ecology Resources* 11, 31-49.

Abadía-Cardoso A, Anderson EC, Pearse DE and Garza JC (2013) Large-scale parentage analysis reveals reproductive patterns and heritability of spawn timing in a hatchery population of steelhead (*Oncorhynchus mykiss*). *Molecular Ecology* 22, 4733-4746.

The co-author Garza JC in these publications directed and supervised the research which forms the basis for the dissertation. The contribution of Abadía-Cardoso A to these publications was intellectually major and included laboratory work, data analysis, and manuscript writing.

General Introduction

The genus *Oncorhynchus* is a monophyletic group of salmonid fishes (Stearley & Smith 1993) that comprises approximately 11 species and about 28 named subspecies (Behnke 2002). In North America, the genus *Oncorhynchus* is divided into two main groups: the Pacific salmon that includes coho (*O. kisutch*), Chinook (*O. tshawytscha*), sockeye (*O. nerka*), chum (*O. keta*), and pink (*O. gorbuscha*) salmon; and the Pacific trout that includes steelhead/rainbow (*O. mykiss* ssp.), cutthroat (*O. clarkii* ssp.), gila (*O. gilae*), apache (*O. apache*), and Mexican golden (*O. chrysogaster*) trout, as well as a diverse complex of taxonomically unclassified trout from the Sierra Madre Occidental (SMO) in Mexico (Behnke 2002; Utter & Allendorf 1994).

Pacific trout taxonomy and classification have been subject to constant debate. Initially, the Pacific trout complex was classified within the genus *Salmo* based on morphological characteristics and life history traits (*e.g.* iteroparity). More recently, other morphological characteristics and molecular techniques were used to reclassify the group into the genus *Oncorhynchus* with the Pacific salmon (Smith & Stearley 1989).

The genus *Oncorhynchus* diverged from the genus *Salmo* around 15-20 million years ago (mya) in the early Miocene (Behnke 1992; Devlin 1993; Wilson & Turner 2009). During the Miocene-Pleistocene, strong geologic activity and climate variability in northwestern North America allowed radiation of the salmonids (Montgomery 2000). By the end of the Miocene the genus had diverged into two distinct lineages: the Pacific salmon, and the Pacific trout, both found in North American drainages (Behnke 1992; Stearley & Smith 1993; Wilson & Turner 2009). The Pacific trout diverged into the rainbow (*O. mykiss*) and cutthroat (*O. clarkii*) lineages during the late Pleistocene

(Behnke 1992; Crespi & Fulton 2004). The current native distribution of *O. mykiss* extends from the Kamchatka Peninsula in northeastern Asia to northern Mexico in North America. However, it has been introduced worldwide and there are now naturalized populations of the species in Europe (Fausch 2007), and in the southern hemisphere in Argentina (Pascual *et al.* 2001), New Zealand (Scott 1978), and many other places. Two phylogenetically distinct lineages within *O. mykiss* have been identified in North America: the inland and the coastal groups, which are roughly separated by the Cascades mountain range (Behnke 1992; Burgner *et al.* 1992; Busby *et al.* 1996). Within these groups, several ecotypes of *O. mykiss* can be distinguished. The nonanadromous freshwater resident form is called rainbow, golden or redband trout and the anadromous form of the species is called steelhead. Steelhead spend several years (up to seven) in freshwater, then migrate to the ocean where they spend up to three years, before coming back to freshwater to spawn. Steelhead also show distinct temporal "runs" or "races" that are defined by the season (spring, summer, fall or winter) of peak river entry and associated reproductive maturity (Busby *et al.* 1996).

These extremely complex life history traits present great difficulty for the assessment and monitoring of the species populations. Moreover, salmonid populations on the West coast of the United States have declined dramatically during the past few decades and many steelhead populations are now protected under the United States Endangered Species Act (NOAA 2006). The most important causes for this decline include habitat loss, habitat degradation, recreational and commercial fishing, and hatchery operations. Introgression by genetically depauperate hatchery rainbow trout that have

been stocked in great numbers in basins containing native steelhead has also been reported as a potential threat to some steelhead populations (Araki *et al.* 2007a; Araki *et al.* 2007b; Clemento *et al.* 2009; Garza & Pearse 2008). But the threats faced by these salmonid species are not exclusive to the United States. Trout inhabiting Northwestern Mexico are likely to go extinct due to these threats without urgent documentation and conservation action. The Mexican trout complex has been recognized as one of the most diverse and least known groups of trout, since there is scarce knowledge about their taxonomic status (Behnke 2002).

The fossil record indicates that trout inhabited Mexico during the Pleistocene. The southernmost record for a fish assigned to the family Salmonidae is in the Lake Chapala, Jalisco, Mexico region near 20° North latitude (Cavender & Miller 1982).

Behnke (1992) suggests that the Gulf of California acted as a refugium for anadromous *O. mykiss* during the Pleistocene glaciations. These trout migrated from the Gulf into rivers of northwestern Mexico, Arizona and New Mexico. The subsequent increase in both ocean and river water temperatures constrained these trout to the high elevation headwaters of different river systems. Long isolation times gave rise to the Gila (*O. gilae*), Apache (*O. apache*), Mexican golden (*O. chrysogaster*), and, presumably, the other SMO trout.

Molecular population genetic analysis has proven to be one of the most effective methods for addressing phylogenetic, ecological and conservation questions and for providing other types of biological inference on fishes. Considerable interest has been shown for decades in the phylogeny and taxonomic status of the rainbow trout complex.

Through the years, multiple molecular methods have been used to understand the genetic identity of subgroups within the complex throughout western North America. For example, allozyme analysis showed the longitudinal separation of the inland and coastal *O. mykiss* lineages (Allendorf 1975), and also, gave insights into the genetic structure along the coast, indicating stronger genetic similarity within geographically proximate populations (Okazaki 1984; Utter *et al.* 1973). Mitochondrial DNA (mtDNA) analysis was the common method used during the 1980s and 1990s to identify genetic structure among natural populations and differentiation from hatchery-raised fish (Nielsen *et al.* 1994a; 1994b; Bagley & Gall 1998; McCusker *et al.* 2000). Analysis using mtDNA further revealed the phylogenetic proximity of the two *O. mykiss* ecotypes (rainbow trout and steelhead), and its divergence from cutthroat trout (Wilson *et al.* 1985; Thomas *et al.* 1986). More recently, microsatellite loci are widely used as a tool for investigating population structure and interactions among different groups of *O. mykiss* with very high resolution (Beacham *et al.* 2000; Aguilar & Garza 2006; Pearse *et al.* 2007; Clemento *et al.* 2009). Additionally, due to its economic importance, other genomic resources have been developed for the species, including expressed sequence tag (EST) databases (http://compbio.dfci.harvard.edu) and linkage maps (Rexroad *et al.* 2008). In turn, these resources allow the development of valuable and powerful genetic tools, such as single nucleotide polymorphism (SNP) markers, that improve our understanding of the biology and evolutionary relationships within the rainbow trout complex, to be able to enact appropriate management strategies for these fishes. A SNP represents a mutation occurring on a DNA sequence when a single nucleotide base differs between

two alleles or individuals. Several characteristics make SNPs extremely powerful. For example, SNPs are the most common type of genetic variation in vertebrates (Wang *et al.* 1998; Smith *et al.* 2005), they can occur within any genomic region, and they have a relative low mutation rate (Brumfield *et al.* 2003). In addition, SNPs offer the potential for high-throughput genotyping at low cost, lower genotyping error rates and ease of standardization between laboratories.

In spite of all that is known about the rainbow trout complex, there are still many questions that remain and that can only be resolved by applying new genomic and molecular resources. An extensive and in depth evaluation of genetic variability in the rainbow trout complex was undertaken to address some of these evolutionary and conservation questions.

In chapter one, I have described a large number of SNPs for *O. mykiss* in order to study ecological interactions, phylogeography, and conservation status (Abadía-Cardoso *et al.* 2011). These molecular genetic tools have proven to have great power, not only for traditional population genetic analysis and phylogenetics, but also for individual identification and for the reconstruction of pedigrees. The last of these allows the tracking of families, and an unprecedented level of monitoring and evaluation of natural and hatchery/aquaculture populations (Anderson & Garza 2006). Chapter two expands the scope of these new molecular genetic tools, by demonstrating the application of a novel individual-based method for large-scale reconstruction of pedigrees in a steelhead population. This work provides a powerful approach for understanding many basic biological traits in a relatively high fecundity species with significant con-

servation concerns, including estimation of variance in reproductive success, migration rates, effective population sizes, life-stage-specific mortality rates, and other population parameters. In this chapter, I was able to elucidate these life history patterns for steelhead populations from two hatchery programs in the Russian River to examine whether assumptions made by resource managers are supported and whether supplementation may be negatively influencing the associated natural populations (Abadía-Cardoso *et al.* 2013). In chapter three, the powerful molecular tools developed were applied to population genetic analysis of the Mexican trout complex, to evaluate population structure and differentiation, and to understand its phylogeographic distribution. In addition, an evaluation of the extent of hybridization and genetic introgression from exotic hatchery rainbow trout into the native trout populations of northwestern Mexico was performed, all to better understand the evolutionary origins of this group and to contribute to the conservation of its important biodiversity.

References

Abadía-Cardoso A, Anderson EC, Pearse DE, Garza JC (2013) Large-scale parentage analysis reveals reproductive patterns and heritability of spawn timing in a hatchery population of steelhead (*Oncorhynchus mykiss*). Molecular Ecology 22, 4733-4746.

Abadía-Cardoso A, Clemento AJ, Garza JC (2011) Discovery and characterization of single nucleotide polymorphisms in steelhead/rainbow trout, *Oncorhynchus mykiss*. Molecular Ecology Resources 11, 31-49.

Aguilar A, Garza JC (2006) A comparison of variability and population structure for major histocompatibility complex and microsatellite loci in California coastal steelhead (*Oncorhynchus mykiss* Walbaum). Molecular Ecology 15, 923-937.

Allendorf FW (1975) Genetic variability in a species possessing extensive gene duplication: Genetic interpretation of duplicate loci and examination of genetic variation in populations of rainbow trout. Ph.D. Dissertation. University of Washington, Seattle.

Anderson EC, Garza JC (2006) The power of single-nucleotide polymorphisms for large-scale parentage inference. Genetics 172, 2567-2582.

Araki H, Cooper B, Blouin MS (2007) Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. Science 318, 100-103.

Araki H, Waples RS, Ardren WR, Cooper B, Blouin MS (2007) Effective population size of steelhead trout: influence of variance in reproductive success, hatchery programs, and genetic compensation between life-history forms. Molecular Ecology 16, 953-966.

Bagley MJ, Gall GAE (1998) Mitochondrial and nuclear DNA sequence variability among populations of rainbow trout (*Oncorhynchus mykiss*). Molecular Ecology 7, 945-961.

Beacham TD, Pollard S, Le KD (2000) Microsatellite DNA population structure and stock identification of steelhead trout (*Oncorhynchus mykiss*) in the Nass and Skeena Rivers in northern British Columbia. Marine Biotechnology 2, 587-600.

Behnke RJ (1992) Native trout of western North America American Fisheries Society Monograph 6, Bethesda.

Behnke RJ (2002) Trout and salmon of North America Chanticleer press, inc., New York.

Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. Trends in Ecology & Evo-

lution, 18, 249-256.

Burgner RL, Light JT, Margolis L, *et al.* (1992) Distribution and origins of steelhead trout *Oncorhynchus mykiss* in offshore waters of the north Pacific Ocean. International North Pacific Fisheries Commission Bulletin 51, 1-92.

Busby PJ, Wainwright TC, Bryant GJ, *et al.* (1996) Status review of West coast steelhead from Washington, Idaho, Oregon, and California. US Department of Commerce.

Cavender TM, Miller RR (1982) *Salmo australis*, a new species of fossil sal-monid from southwestern Mexico. In: Contributions from the museum of paleontology (ed. Smith GR), pp. 1-17. The University of Michigan.

Clemento AJ, Anderson EC, Boughton D, Girman D, Garza JC (2009) Population genetic structure and ancestry of *Oncorhynchus mykiss* populations above and below dams in south-central California. Conservation Genetics 10, 1321-1336.

Crespi BJ, Fulton MJ (2004) Molecular systematics of Salmonidae: combined nuclear data yields a robust phylogeny. Molecular Phylogenetics and Evolution 31, 658-679.

Devlin RH (1993) Sequence of sockeye-salmon type-1 and type- 2 growth-hormone genes and the relationship of rainbow-trout with Atlantic and Pacific salmon. Canadian Journal of Fisheries and Aquatic Sciences 50, 1738-1748.

Fausch KD (2007) Introduction, establishment and effects of non-native sal-monids: considering the risk of rainbow trout invasion in the United Kingdom. Journal of Fish Biology 71, 1-32.

Garza JC, Pearse DE (2008) Population genetic structure of *Oncorhynchus mykiss* in the California Central Valley. In: Final report for California Department of Fish and Game Contract No. PO485303, p. 54. University of California, Santa Cruz and NOAA, Southwest Fisheries Science Center.

Gene Index. http://compbio.dfci.harvard.edu Dana-Farber Cancer Institute and Harvard School of Public Health.

McCusker MR, Parkinson E, Taylor EB (2000) Mitochondrial DNA variation in rainbow trout (*Oncorhynchus mykiss*) across its native range: testing biogeographical hypotheses and their relevance to conservation. Molecular Ecology 9, 2089-2108.

Montgomery DR (2000) Coevolution of the Pacific salmon and Pacific Rim topography. Geology 28, 1107-1110.

Nielsen JL, Gan C, Thomas WK (1994) Differences in genetic diversity for mitochondrial DNA between hatchery and wild populations of *Oncorhynchus*. Canadian Journal of Fisheries and Aquatic Sciences 51, 290-297.

Nielsen, JL, Gan C, Wright JM, Morris DB, *et al.* (1994) Biogeographic distributions of mitochondrial and nuclear markers for southern steelhead. Molecular Marine Biology and Biotechnology 3, 281-293.

NOAA (2006) Endangered and threatened species: final listing determinations for 10 distinct population segments of West Coast steelhead (ed. NOAA), pp. 833-862. US Federal Register.

Okazaki T (1984) Genetic divergence and its zoogeographic implications in closely related species *Salmo gairdneri* and *Salmo mykiss*. Japanese Journal of Ichthyology 31, 297-311.

Pascual M, Bentzen P, Rossi CR, *et al.* (2001) First documented case of anadromy in a population of introduced rainbow trout in Patagonia, Argentina. Transactions of the American Fisheries Society 130, 53-67.

Pearse DE, Donohoe CJ, Garza JC (2007) Population genetics of steelhead (*Oncorhynchus mykiss*) in the Klamath River. Environmental Biology of Fishes 80, 377-387.

Rexroad III CE, Palti Y, Gahr SA, Vallejo RL (2008) A second generation genetic map for rainbow trout (*Oncorhynchus mykiss*). BMC Genetics 9, 74.

Scott D (1978) The origins of rainbow trout, *Salmo gairdneri* Richardson, in New Zealand. California Fish and Game 64, 200-209.

Smith CT, Elfstrom CM, Seeb LW, Seeb JE (2005) Use of sequence data from rainbow trout and Atlantic salmon for SNP detection in Pacific salmon. Molecular Ecology, 14, 4193-4203.

Smith GR, Stearley RF (1989) The classification and scientific names of rainbow and cutthroat trouts. Fisheries 14, 4-10.

Stearley RF, Smith GR (1993) Phylogeny of the Pacific trout and salmon (*Oncorhynchus*) and genera of the family Salmonidae. Transactions of the American Fisheries Society 122, 1-33.

Thomas WK, Withler RE, Beckenbach AT (1986) Mitochondrial DNA analysis of Pacific salmonid evolution. Canadian Journal of Zoology 64, 1058-1064.

Utter FM, Allendorf FW (1994) Phylogenetic-relationships among species of *Oncorhynchus*: a consensus view. Conservation Biology 8, 864-867.

Utter FM, Allendor FW, Hodgins HO (1973) Genetic variability and relationships in Pacific salmon and related trout based on protein variations. Systematic Zoology 22, 257-270.

Wang DG, Fan JB, Siao CJ, *et al.* (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science, 280, 1077-1082.

Wilson GM, Thomas WK, Beckenbach AT (1985) Intra-specific and inter-specific mitochondrial DNA sequence divergence in *Salmo*: Rainbow, steelhead, and cutthroat trouts. Canadian Journal of Zoology 63, 2088-2094.

Wilson WD, Turner TF (2009) Phylogenetic analysis of the Pacific cutthroat trout (*Oncorhynchus clarki* ssp.: Salmonidae) based on partial mtDNA ND4 sequences: A closer look at the highly fragmented inland species. Molecular Phylogenetics and Evolution 52, 406-415.

Chapter 1

Discovery and characterization of single nucleotide polymorphisms (SNPs) in steelhead/rainbow trout *Oncorhynchus mykiss*

Abstract

Single nucleotide polymorphisms (SNPs) have several advantages over other genetic markers, including lower mutation and genotyping error rates, ease of inter-laboratory standardization, and the prospect of high-throughput, low-cost genotyping. Nevertheless their development and use has only recently moved beyond model organisms to groups such as salmonid fishes. *Oncorhynchus mykiss* is a salmonid native to the North Pacific rim that has now been introduced throughout the world for fisheries and aquaculture. The anadromous form of the species is known as steelhead. Native steelhead populations on the west coast of the United States have declined and many now have protected status. The non-anadromous, or resident, form of the species is termed rainbow, redband or golden trout. Additional life history and morphological variation, and interactions between the forms, make the species challenging to study, monitor and evaluate.

Here I describe the discovery, characterization and assay development for 139 SNP loci in steelhead/rainbow trout. I used EST sequences from existing genomic databases to design primers for 480 genes. Sanger-sequencing products from these genes provided 130KB of consensus sequence in which variation was surveyed for 22 individuals from steelhead, rainbow and redband trout groups. The resulting TaqMan assays were surveyed in 11 natural-origin steelhead populations, four *O. mykiss* hatchery strains, and two introduced *O. mykiss* populations in the southern hemisphere, where they had a mean minor allele frequency of 0.112-0.258 and observed heterozygosity of

0.0001-0.342. Mean F_{ST} was 0.206. All 139, along with 28 assays previously developed by other laboratories, where screened to select a panel of 96 highly informative SNPs. The selection of assays was based on their utility for parentage inference in four steelhead populations in California and their ability to distinguish individuals from several California populations. The development of SNPs for *O. mykiss* will help to provide highly valuable genetic tools for individual and stock identification, pedigree reconstruction, phylogeography, and ecological investigation.

1.1 Introduction

The development of highly informative molecular markers is an important first step in the investigation of population, ecological, evolutionary and conservation genetic questions. Several types of molecular markers have been widely used since the development of the polymerase chain reaction (PCR), including randomly amplified polymorphic DNA (RAPDs), amplified fragment length polymorphisms (AFLPs), mitochondrial DNA sequences and variable number of tandem repeat markers, such as microsatellites and minisatellites. More recently single nucleotide polymorphisms (SNPs) have begun to see use in population genetics, although primarily for model organisms. SNPs are nucleotide variants found at particular genomic locations and are normally bi-allelic (Vignal *et al.* 2002). SNPs have several advantages over other markers, including that they are the most abundant polymorphisms in vertebrate genomes, with an approximate density of 10^{-3} SNPs per base pair (Smith *et al.* 2005; Wang *et al.* 1998), they

are found in both coding and noncoding regions (Brumfield *et al.* 2003), and they have a lower mutation rate (Brumfield *et al.* 2003), which is an important source of error in many applications. The use of SNP markers with humans and other model organisms is extensive and has focused on genetic mapping, disease diagnosis, toxicology and pharmacogenomics (McCarthy & Hilfiker 2000; Sachidanandam *et al.* 2001; Wang *et al.* 1998). Conversely, in non-model organisms, such as salmonid fishes, the use of SNP markers is quite recent and has focused more on population identification and ecological genetic questions (Narum *et al.* 2008).

Oncorhynchus mykiss is a salmonid species native to the North Pacific rim. Its current native distribution extends from the Kamchatka Peninsula in northeastern Asia to northern Mexico in North America. However, it has been introduced throughout the world for recreational fisheries and aquaculture, and there are now naturalized populations of the species in the southern hemisphere (Pascual *et al.* 2001) and in Europe (Fausch 2007). Two widespread and phylogenetically distinct lineages of *O. mykiss* have been identified in North America and they correspond roughly to inland and coastal groups separated by the Cascades mountain range (Burgner *et al.* 1992; Busby *et al.* 1996), although the full phylogenetic picture is more complicated (McCusker *et al.*, 2000). In addition, many ecotypes and life history strategies are present in the species. Generally, the anadromous form of the species is termed steelhead and the nonanadromous, freshwater form rainbow, golden or redband trout. Steelhead spend from one to seven years in fresh water, then migrate to the ocean where they spend from one to three years before returning to fresh water to spawn. However, life history

strategy in *O. mykiss* is governed by a complex mix of environmental and heritable factors, such that a single interbreeding population can contain individuals expressing nearly every possible combination of years in fresh and salt water (Shapovalov & Taft 1954). There are also several ecotypes of steelhead that can coexist as distinct temporal "runs" or "races" that are defined by the season (spring, summer, fall or winter) of peak river entry and associated reproductive maturity (Busby *et al.* 1996).

This life history complexity makes monitoring and evaluation of the species, and its multitude of managed populations and stocks, difficult. Such assessment has become increasingly important, since salmonid populations on the west coast of the United States have declined dramatically during the past few decades and many steelhead populations are now protected under the United States Endangered Species Act (ESA; NOAA, 2006). The most important causes for this decline include habitat loss, habitat degradation, recreational harvest and hatchery operations. In addition, genetically depauperate hatchery rainbow trout have been stocked in great numbers in basins containing native steelhead. Introgression by these trout has been reported and may pose a substantial threat to at least some steelhead populations (Clemento *et al.* 2009; Garza & Pearse 2008).

One of the most important methods for monitoring the effects of such threats on fish populations, and for providing other types of biological inference about them, is the use of molecular population genetic analysis. Microsatellite loci have seen widespread use with *O. mykiss* and have proven powerful in studying population structure and interactions among different groups (Aguilar & Garza 2006; Beacham *et al.* 2000;

Clemento *et al.* 2009; Narum *et al.* 2004; Pearse *et al.* 2007). Fortunately, due primarily to the importance of *O. mykiss* in aquaculture, many additional genomic resources have been developed for the species, including expressed sequence tag (ESTs) databases and linkage maps (Rexroad *et al.* 2008).

These resources are allowing more detailed analyses of ecological and conservation genetic questions than previously possible (*e.g.* Martínez *et al.* 2011). They also allow the identification and development of SNP markers for salmonid species that can be surveyed on a large scale (Castaño-Sánchez *et al.* 2009; Smith *et al.* 2005). Such markers will allow large-scale monitoring and will further elucidate some of the pressing questions regarding *O. mykiss* ecology and life history evolution, through both traditional population genetic analyses and through large-scale parentage inference (Anderson & Garza 2006), particularly with the advent of high-throughput genotyping methods.

In this study I describe the discovery, characterization and development of assays for a large number (139) of SNP loci for steelhead/rainbow trout. I exploited EST databases to design nearly 500 primer sets for functional genome regions. PCR products resulting from these genes, which include both exonic and intronic regions, were then sequenced in an ascertainment panel of 22 fish designed to simultaneously represent some of the phylogenetic diversity of the species and to provide polymorphic markers for focal populations in California. Such "balanced" ascertainment is intended to reduce the bias against polymorphism in other populations and lineages of a species when only particular groups are used in marker discovery (Clark *et al.* 2005). I also in-

cluded two individuals of a sister species, *O. clarkii*, in order to identify species diagnostic markers. These SNP markers represent a valuable resource for studying ecological interactions, phylogeography, and conservation status, as well as for pedigree reconstruction, individual and genetic stock identification and, eventually, for linkage mapping.

1.2 Methods

1.2.1 *Ascertainment panel*

Individuals from multiple populations and lineages of *O. mykiss* were chosen for the ascertainment panel. A total of 22 fish from five distinct steelhead populations or rainbow trout strains were included: 10 anadromous adult steelhead from Scott Creek, four anadromous adult steelhead from the Middle Fork Eel River summer run, two redband trout (*Oncorhynchus mykiss newberrii*) from the Upper Klamath River basin, and six hatchery rainbow trout raised at Fillmore Hatchery on the Santa Clara River near Los Angeles, CA. Three of these trout were from either the Virginia or Wyoming strains and three were from the Mt. Whitney Strain (Busack & Gall 1980). In addition, two coastal cutthroat trout (*O. clarkii clarkii*) from Little River, Humbolt County, CA were included in the ascertainment panel, in order to detect and avoid designing assays for polymorphisms that might be due to past hybridization between steelhead and cutthroat trout (Young *et al.* 2001), and to identify candidate markers for species diagnostic.

1.2.2 *Genetic analysis*

Tissue samples were digested with proteinase K, followed by DNA extraction with a semi-automated membrane-based system (DNeasy 96 Tissue Kit, QIAGEN Inc.) on a QIAGEN BioRobot 3000. All of these samples had been previously genotyped with microsatellites, so that DNA quality was known to be high. Purified DNA was diluted 1:20 in ddH_20 for PCR.

A total of 480 *O. mykiss* expressed sequence tags (ESTs) were selected using a random number generator from the rainbow trout "Gene Index" (RtGI) online database hosted at the Dana-Farber Cancer Institute and Harvard School of Public Health (http://compbio.dfci.harvard.edu; accessed on December 8, 2006). Primers were designed using the program primer3 v.0.4.0 (Rozen & Skaletsky 2000) for each of these loci. PCR amplifications were conducted using the following parameters: 0.041 U AmpliTaq DNA polymerase (Applied Biosystems Inc.), 1.5μL PCR buffer (Applied Biosystems Inc.), 0.9mM $MgCl_2$, 0.5mM dNTPs, 5μmol of each primer and 14μL of DNA template. Thermal cycling conditions employed a "touchdown" protocol and were as follows: an initial denaturation of 3 min at 94°C, then 2 min at 63°C, and 1 min at 72°C, followed by [94°C for 30s, 60°C for 30s, 72°C for 1 min] x 12 (-1°C/cycle), [94°C for 30s, 48°C for 30s, 72°C for 1 min] x 11; [94°C for 30s, 48°C for 30s, 72°C for 1 min (+ 10s/cycle)] x 9, and finally 5 min at 72°C. PCR products were surveyed by gel electrophoresis in 2% agarose. PCR products that exhibited a single robust band were purified using an Exo-Sap protocol (USB Inc.): 5mL of PCR product, 0.15mL of

Exonuclease I (20U/mL), 1mL of shrimp alkaline phosphatase (1U/mL), 0.5mL of 10x buffer and 3.36mL of deionized water were incubated at 37°C for 60 min and then 80°C for 20 min with a cool down to 4°C. Clean products were then Sanger-sequenced on both the forward and reverse strands using the BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems Inc.). Sequencing reaction products were purified using 6% Sephadex columns and visualized by capillary electrophoresis on a 3730 DNA Analyzer (Applied Biosystems Inc.).

All sequences from each locus were aligned and assembled into contigs using Sequencher 4.9 (Gene Codes Corporation). Where the alignments indicated a polymorphism, the chromatograms were visually examined for verification. To consider a polymorphism for development as a SNP assay, I used the criterion that all three genotypes (the homozygotes for both alleles and the heterozygote) for that site must have been observed at least once in the ascertainment panel. No distinction was made with respect to the population or strain in which the genotypes were found. This ascertainment criterion was employed to reduce the identification of sequencing artifacts as SNPs and to select the nucleotide sites that had the highest probability of being sufficiently polymorphic for downstream applications. A BLAST search was also performed on each consensus sequence to determine if the EST corresponded to an identified gene and to ensure that each SNP marker would represent a novel assay in an independent gene. I chose one potential SNP for each EST analyzed in order to reduce the probability of markers in linkage disequilibrium. The site with the highest minor allele frequency in the ascertainment sample that also met the assay design criteria (*e.g.* more than 25bp

from the end of the sequence, no adjacent polymorphism) was chosen for assay design. The original ESTs and the BLAST results were also used to identify the variability patterns of the SNPs chosen, such as the region on the gene (intronic or exonic) and the position on the triplet (synonymous or non-synonymous).

1.2.3 *SNP assay development and validation*

Consensus sequences, with the selected nucleotide sites indicated, were submitted for design of 5' nuclease allelic discrimination, or TaqMan, assays (Applied Biosystems Inc.). When it was not possible to design an assay for a selected site and another nucleotide in the consensus sequence met both the ascertainment and design criteria, a second attempt was made to design an assay for that locus.

SNP assays were validated by genotyping a total of 376 fish from 11 natural-origin steelhead populations, four *O. mykiss* hatchery strains, and two introduced *O. mykiss* populations in the southern hemisphere.

SNP genotyping was carried out in 96.96 Dynamic Genotyping Arrays on an EP1 Genotyping System (Fluidigm Corporation), which uses nanofluidic circuitry to simultaneously interrogate up to 96 loci in 96 individuals.

1.2.4 *Statistical analysis*

Deviations from Hardy-Weinberg and gametic phase (linkage) equilibrium were evaluated with GENEPOP 4.0 (Rousset 2008). Observed and expected heterozygosity (Nei 1978), the fixation index F_{ST} (Weir & Cockerham 1984), and allele frequencies

were estimated using GENETIX 4.05 (Belkhir *et al.* 1996-2004).

1.2.5 *Steelhead SNP panel development*

All 139 SNPs described in this study, along with 28 assays previously developed by other laboratories (Aguilar & Garza 2008; Campbell *et al.* 2009, CRITFC - N. Campbell unpubl.; WSU - J. DeKoning unpubl.), where screened to select a panel of 95 highly informative SNPs. The selection of assays was based on their utility for parentage inference in four steelhead populations in California and their ability to distinguish individuals from several California populations.

In order to identify the patterns of variability of the SNPs in the panel, as well as the rest of the SNPs developed in this study, I performed an alignment exercise in which the consensus sequences were aligned with the RtGI original ESTs -from which the primers were designed in the sequencing effort. That alignment helped identifying if the SNPs were in intronic or other regions of the gene (coding DNA sequences (CDS) and un-translated regions (UTR)). Then, the results from the above BLAST search were used to confirm these regions as well as to obtain the protein translation, so as to identify if the SNP corresponded to a synonymous or a non-synonymous mutation.

1.3 Results

Of the 480 primer pairs designed from *O. mykiss* ESTs, 264 produced a single-sized PCR product in most or all fish in the ascertainment panel. Of these 264 ESTs, 236 yielded sequence from one or more individuals. All PCR products were sequenced,

even if a band was not visible for every individual on an agarose gel. A mean of 18 (range 1-22) individuals produced sequence for each locus, and most of these resulted in broadly or completely overlapping forward and reverse sequences. Since EST sequences are derived from mRNA and therefore lack intronic regions, many of the PCR products were larger than the predicted size, and several of them did not have overlapping forward and reverse strand sequences. None of the ESTs were identified as coming from the same gene in a BLAST search, nor did they match any published SNP assays for *O. mykiss*.

More than 2.3 MB of genomic sequence was produced and aligned (Table 1.1), or 4.6 MB when both strands were considered separately, and a composite consensus sequence of 130KB (mean 551bp/locus) was used for discovery and the determination of density. To account for the lack of sequence for all individuals in all sequences, and the consequent decrease in probability of finding variability, I calculated a consensus length weighted by the number of individuals for which sequence was obtained. The weighted consensus sequence was 120KB (mean 513bp/locus). In other words, 92.3% (120KB/130KB) of the entire consensus sequence from these 236 loci was obtained for all 22 individuals in the ascertainment panel. The density of all nucleotide sites with apparent substitutions was 0.0111, or one every 111 bp. When weighted by the number of fish for which sequenced was obtained, the density of substitutions was 0.0122 or one every 122 bp.

A total of 175 sequences were submitted for assay design. In addition, one sequence (GHPROM1) with a SNP identified in a previous effort (Aguilar & Garza 2008) was submitted for design. Of those, 167 yielded designs suitable for assay man-

Table 1.1: Summary of EST sequencing effort.

	Total	Mean [Range] per locus
Loci sequenced	236	
Base-pairs sequenced	2,322,269	
Length of Consensus sequence (base pairs)	130,025	550.95 [109-1417]
Weighted consensus (base pairs)	119,969	512.69
Number of Substitutions	1,366	5.84 [0-21]
Number of SNPs (all three genotypes)	506	2.16 [0-10]
Loci with no variable sites	10	
Indels	182	
Transitions (A-G or C-T)	676	
Transversions (A-C or G-C or A-T or G-T)	681	
Possible duplicated genes	14	
Tri-nucleotides	9	
Total number of SNPs + indels	1,548	
Density of substitutions in consensus sequences		0.0111
Density of substitutions in weighted consensus sequences		0.0122

ufacture. From these 167, I then eliminated 28 because of problems with genotyping calling or because the assay was not interrogating a single Mendelian locus (all apparent homozygotes or heterozygotes).

This elimination process left 139 SNP assays for further validation and characterization. A list of these assays, with primer/probe information and with the variable base indicated, is found in Table 1.2. To evaluate the utility of these loci in different parts of the species' geographic range, and for both natural populations and hatchery/aquaculture rainbow trout, I genotyped all 139 loci in 17 steelhead populations or rainbow trout strains (Table 1.3). Several locI was not in Hardy-Weinberg equilibrium for some populations or strains, but only four loci deviated from equilibrium in more than one group and no locus deviated in more than three populations or strains. Very little linkage disequilibrium between markers was found. Three markers (Omy_114448-87, Omy_121006-131 and Omy_127236-583) were in complete disequilibrium, in spite of

the fact that they were designed from unique ESTs, but aside from those three, only eight pairs of markers (out of a total of 9005 pairs), were in significant linkage disequilibrium ($p < 0.001$; 53 more pairs if $p < 0.01$), which is similar to the number expected by chance alone.

Table 1.2: SNP type, forward and reverse primers (5'-3'), TaqMan probes and dye, length of consensus sequences, GenBank accession numbers and dbSNP accession numbers for the 139 SNP loci. GB: GenBank Number; db: dbSNP Number.

Assay name	Assay targets	Primers (5'-3')	Probes (5'-3')	Cons. length	GeneBank No. dbSNP No.
OMGH1PROM1 -SNP1	A/T	fwd: TCAAACTGCATTTGATGGAAACAAACAT rev: AGGACAATTCTAAGTGACCTCAAACTG	VIC: TAGTGTTCACTGACTTCA FAM: TAGTGTACACTGACTTCA	n/a	GB: n/a db: n/a
Omy_95318-147	C/T	fwd: CGTGTTCTATTTGAAGGCTGTTAAAGG rev: TCCTGAACTTAAACCTGCCGTTTT	VIC: TTGGTTCTGGATATTAT FAM: CTTGGTTCTGAATATTAT	377	GB: HR504810 db: 275517390
Omy_95442-108	G/A	fwd: CCCTGATTATATAGGGAGCTTTCCACTT rev: CTTCGGCTCTCGCCAAGT	VIC: TTTCCCACCCAGCATT FAM: CTTTCCCACTCAGCATT	255	GB: HR504811 db: 275517391
Omy_95489-423	T/C	fwd: TGAGTCCAGTAAAATCCAATCAATAATCATGT rev: ACTAGAGCACTGATAGGCTGTCA	VIC: CTGCCACTAACATAC FAM: CTGCCACTGACATAC	595	GB: HR504812 db: 275517392
Omy_96158-277	T/G	fwd: TTGTGACGGGATCCTTCATTGAG rev: CACCTCGATCTCCTTTCGGTAAAA	VIC: AAATACGACCCAACAATA FAM: AAATACGACCCCACAATA	336	GB: HR504813 db: 275517393
Omy_96222-125	T/C	fwd: GTAAGGAACTAATTGGCGCAACATT rev: CAGTTTGTCTAACACCCAGGCATAT	VIC: AACTACAACTGTAGCTAATT FAM: CAACTGTGGCTAATT	615	GB: HR504814 db: 275517394
Omy_96529-231	C/T	fwd: GCGGTCCACAACCTTCTATCC rev: GCCACGGCAAGGTTAAGG	VIC: ATTTCACATAGTGGGTCTG FAM: ATTTCACATAATGGGTCTG	305	GB: HR504815 db: 275517395
Omy_96899-148	T/G	fwd: CCGGACTACAGGGCTCTGA rev: GTGACCCTCCCAGCTTCTG	VIC: CAGGCCTTAGTGCAGC FAM: AGGCCTTCGTGCAGC	568	GB: HR504816 db: 275517396
Omy_97077-73	T/A	fwd: GTGTAAACAAAATGACTCTGGGATTCAG rev: AGAAGTGGCAATGGTGTGAAGTAT	VIC: TGGTGCAATAGAAATA FAM: CATGGTGCAATAGTAATA	295	GB: HR504817 db: 275517397
Omy_97660-230	C/G	fwd: TCAGTTATGTGTAATCTCATTACCTCTCCAA rev: AACAGAAAAGGTCTCAATGTATTTTTTGCA	VIC: ACGTAACTTGTAGCGTTTT FAM: ACGTAACTTGTACCGTTTT	461	GB: HR504818 db: 275517398
Omy_97865-196	A/G	fwd: TCCAGACTTCTGGTTTGTTCCATT rev: CCAGCCCCTATATTCACAATTAAGTGT	VIC: ATGAGCTTGTTAATTAAT FAM: AGCTTGTCAATTAAT	299	GB: HR504819 db: 275517399
Omy_97954-618	C/T	fwd: GCTCTGCTTCCTCGGCAAATA rev: CACAATTGGTTTTTGCACAAAAGTAAAGTATT	VIC: CAACGCTTACCGGTGTGT FAM: CAACGCTTACCAGTGTGT	871	GB: HR504820 db: 275517400
Omy_98188-405	T/C	fwd: CACAGTTGCAAGGTAGAGGCTTATA rev: GCTGAAAGATTAATCCCAGACTGTAGATT	VIC: CTCTCATAAGTCTATCCTCC FAM: CTCATAAGTCTGTCCTCC	425	GB: HR504821 db: 275517401
Omy_98409-549	A/G	fwd: CGCCTTCCTCAGTATGACATATGTA rev: AGGATTTCAGGAAAACCAGGGAATT	VIC: ATTTTGCAACTCTACTTTC FAM: TTGCAACCCTACTTTC	1077	GB: HR504822 db: 275517402
Omy_98683-165	A/C	fwd: GCCATTGCCAGAGAATTTGGTTAA rev: AACACACGCACCATCTTAAAGC	VIC: AGCCAGATACATATTTGT FAM: CCAGATACAGATTTGT	897	GB: HR504823 db: 275517403
Omy_99300-202	T/A	fwd: CAGTTTGACCCGATGGTGTGA rev: GATTATGGCGTGGCCTTTTGG	VIC: TCAGGCATGAGAGAAA FAM: ATCAGGCATGTGAGAAA	386	GB: HR504824 db: 275517404
Omy_100771-63	T/A	fwd: CATTTAAGGAGGTGGGTTTGTGAAA rev: AGTTTGCTGCCACTTGACAGTATT	VIC: AAAGAGCTAGAAATACTG FAM: AAAGAGCTAGAATTACTG	399	GB: HR504825 db: 275517405
Omy_100974-386	T/C	fwd: ACATGCAAATTAAGTGTTTGTTTTTAAAAATCGAA rev: CGACTTCATCCTTTTCATGTAGTGAGT	VIC: CACAGTATTATCAAGATTTT FAM: CAGTATTATCGAGATTTT	471	GB: HR504826 db: 275517406
Omy_101119-554	A/G	fwd: GGTGGCTTGTTTTCTCCCTGTTT rev: GCTCCTACCAGCTGAACAGA	VIC: CATGGACATGATGTTACC FAM: ATGGACATGACGTTACC	1110	GB: HR504827 db: 275517407
Omy_101341-188	T/C	fwd: CTGGAGATAGAAAATATCACACAGAACAGT rev: CCTCATTTGATGCATGATTCTGTTGT	VIC: TGATATACTGCAGGTTCC FAM: ATATACTGCGGGTTCC	668	GB: HR504828 db: 275517408
Omy_101554-306	T/C	fwd: GCCTGTATTTCTCCTGTATGTGCAT rev: TCAACTTTTGCAAACTTTTTTATTCTTTGTCATTT	VIC: TGCTTCTCACATTTTTA FAM: TGCTTCTCACGTTTTTA	411	GB: HR504829 db: 275517409
Omy_101704-329	A/C	fwd: TGTGTGTTTAACTGACAGAGATGCT rev: GGAGCAGGAGCTCAAGGA	VIC: CACCTCCTCTCGGCTGT FAM: CTCCTCGCGGCTGT	591	GB: HR504830 db: 275517410
Omy_101770-410	T/C	fwd: GTTTTCTATGAGCAGGAGAGGGTTAA rev: CTTAGAAAGTACTTCTTTAAATATCAAATGCATTCAGT	VIC: CCTGTCTTTCAAAACTAA FAM: CTGTCTTTCAGAACTAA	795	GB: HR504831 db: 275517411
Omy_101832-195	A/C	fwd: TGGCTCTGGACCTGTTGAGA rev: CGTCACAGCTATTTTAGGCGTAGT	VIC: TGTAGTCTTTCAGAGTAGTATG FAM: TAGTCTTTCAGAGGAGTATG	611	GB: HR504832 db: 275517412
Omy_101993-189	A/T	fwd: ACAAAACACAGTGGAATTACAATTAACGTT rev: GGAAGTTAAATTTCGCTTCGTCAGAA	VIC: CTTGATTTGCAGCTTGTCAA FAM: TGATTTGCAGCATGTCAA	782	GB: HR504833 db: 275517413
Omy_102213-204	T/G	fwd: AGATGTTAACTACATTCCATGACAATGATTGA rev: GAGTATCTCATTCGCAACACTATGGT	VIC: CTAAAAACCCATTAATTCAAT FAM: AAAACCCATTCATTCAAT	640	GB: HR504834 db: 275517414
Omy_102420-634	T/G	fwd: GGTCGTAGTACACACCTGAGTAAAT	VIC: CCTAAAGCGCTTATCTTAA	732	GB: HR504835

Table 1.2 Continued

Assay name	Assay targets	Primers (5'-3')	Probes (5'-3')	Cons. length	GeneBank No. dbSNP No.
		rev: CACGACACATGCCAGTAGACT	FAM: CTAAAGCGCTTCTCTTAA		db: 275517415
Omy_102457-423	T/G	fwd: CGATGAGTCAAGATAGTCGCTACT	VIC: CCCCCAAAAATGTC	584	GB: HR504836
		rev: GGCGTATGGAATTTAGTAGACTAGATTTTCA	FAM: CCCCCAAAATGTC		db: 275517416
Omy_102505-102	A/G	fwd: CTGCAAACTGACATGGTAGCAAAA	VIC: AACAGGATGTTTTTGC	150	GB: HR504837
		rev: TGCTTGCTTTTTAAAAACAATCTCCCA	FAM: CAGGATGCTTTTGC		db: 275517417
Omy_102510-682	T/G	fwd: AAGATCAGTGTGGCATCAATGTCA	VIC: TTGTCCTCAATATTCAC	732	GB: HR504838
		rev: TCGTGCCTGGATGTAAGTTAACTG	FAM: TTGTCCTCACTATTCAC		db: 275517418
Omy_102867-443	T/G	fwd: CATTTGTTTAATTTGATTTGGCACAACTTCA	VIC: TTTGGGTACATAATTTTT	443	GB: HR504839
		rev: CCCTAGTTCTGTAACACAAGACGTAA	FAM: TGGGTACATCATTTTT		db: 275517419
Omy_103350-395	A/C	fwd: CGCGTGTTGAACCTAGAATGAC	VIC: AGAACCAGGAAATTAACTAC	471	GB: HR504840
		rev: GGAAAATTCCTGCCAATGACACATG	FAM: CCAGGAAATGAACTAC		db: 275517420
Omy_103577-379	T/A	fwd: GGAGTGATCCAAGGTTATGTACCAA	VIC: AAGTGTGCACACGTTCA	759	GB: HR504841
		rev: CCAGCAATTTCCTTTCGAATCATTGA	FAM: AAGTGTGCACTCGTTCA		db: 275517421
Omy_103705-558	T/C	fwd: CTCCAATCGCAAATACCCAGACT	VIC: AGACTTACCCAGAGTGAGAG	658	GB: HR504842
		rev: CGCAGGAGACGGATGCC	FAM: ACTTACCCAGGGTGAGAG		db: 275517422
Omy_103713-53	T/G	fwd: TCATGAGTGAAGCGCACAGAA	VIC: AGGTTACTGGAGAAATCT	423	GB: HR504843
		rev: CTTTAGTAGGAGGTTGTAACCAAGTCA	FAM: ACTGGCGAAATCT		db: 275517423
Omy_104519-624	T/C	fwd: CGTGTGAGTTTGCGGTAAAGAC	VIC: CAGCAGGATACATCCGACT	1061	GB: HR504844
		rev: TGACGAGTCCGTCTTATCATCCT	FAM: AGCAGGATACGTCCGACT		db: 275517424
Omy_104569-114	A/C	fwd: CCGAGGCCGACGTGATC	VIC: CGCCACTCCGACGCC	565	GB: HR504845
		rev: GCGCCTCGCTCATCATCA	FAM: CCACGCCGACGCC		db: 275517425
Omy_105075-162	T/G	fwd: GGAGAAGGACAAGGACATTGGTAAT	VIC: CTTTCTCTCCTACTTTCC	443	GB: HR504846
		rev: AAAGCAGACCACACCATACTTCTC	FAM: CTTTCTCTCCTCCTTTCC		db: 275517426
Omy_105105-448	C/T	fwd: CAATTTGCAAGCAGGGAAAGGTTAT	VIC: AAGGAGAATGCATAATC	810	GB: HR504847
		rev: GTGATGGGCTGCAATTGCTT	FAM: TGAAAGGAGAATACATAATC		db: 275517427
Omy_105115-367	C/G	fwd: GCTCCCTCCGAAGAAATCTCA	VIC: CATGCTGGAGCGCAAT	401	GB: HR504848
		rev: CATACTCGTCAATCACCCAAGCT	FAM: CATGCTGGACCGCAAT		db: 275517428
Omy_105235-713	C/T	fwd: AGGCCATAAAATCAGGCATTAGGAT	VIC: AGAGAGTCAATCGTTGCAAA	788	GB: HR504849
		rev: TGGGCTCTGCAAAGACAAGA	FAM: AGAGAGTCAATCATTGCAAA		db: 275517429
Omy_105385-406	T/C	fwd: ACCTACCCTCACCTGAACTTCA	VIC: CTTGGAACCATTGCTAC	691	GB: HR504850
		rev: CGCTCTTCTGGGCGTATCG	FAM: TTGGAACCGTTGCTAC		db: 275517430
Omy_105386-347	A/C	fwd: CCAGGAAATCGTCAGCTCTATTTAATACAT	VIC: ACATTTCAACTCAATTAATAATTA	438	GB: HR504851
		rev: GAAACCTCCTTCAACCTCTGGATAA	FAM: TACATTTCAACTCAATGAATAATTA		db: 275517431
Omy_105401-363	A/G	fwd: GGCACCCTCATTCACACATACTAT	VIC: CCAAGTACCCTAGGTTGG	419	GB: HR504852
		rev: GTCTTCTCAAATAACCCCTGTGGAT	FAM: CAAGTACCCCAGGTTGG		db: 275517432
Omy_105407-74	T/G	fwd: GGATGGCTTGGAATGTGCAA	VIC: CTCTTTGCGTTTAGTCCTA	472	GB: HR504853
		rev: GCGGATGTACACAAAATACACTCAA	FAM: TCTTTGCGTTTCGTCCTA		db: 275517433
Omy_105714-265	C/T	fwd: CCACTCAGTGCAAGCATGGA	VIC: CTGTTGTTTGAGGTTCAG	476	GB: HR504854
		rev: GCTTTCAATCCTTGGCTCCAATATC	FAM: TGTTGTTTGAGATTCAG		db: 275517434
Omy_105897-101	T/A	fwd: GAACCAATACACAATGCCAAGGATT	VIC: TCTCTCCACAGTTCTC	382	GB: HR504855
		rev: GCTAGGGCTGCTATCTTTGTGATG	FAM: TCTCTCCACTGTTCTC		db: 275517435
Omy_106172-332	T/G	fwd: CCACTTTGTTACTAAATGTTCCCATGAC	VIC: ATGAACAGAATGTAATCTAG	467	GB: HR504856
		rev: ACATTCCAAAGACTGTCACATTCCA	FAM: TGAACAGAATGTCATCTAG		db: 275517436
Omy_106313-445	T/G	fwd: CCAACTGTTGTGTCTTGATTTGTGA	VIC: TTGATTTTTCCAAACCATGTGTG	729	GB: HR504857
		rev: GTTCTGTGTCTGAAGTCCATTGGT	FAM: TTGATTTTTCCAAACCCTGTGTG		db: 275517437
Omy_106560-58	C/T	fwd: CCACCCAGCCATCAACGA	VIC: CTCAGAGCGCAGGCC	387	GB: HR504858
		rev: CGTTCTTTCCCAGCGAGTGA	FAM: CTCAGAGCACAGGCC		db: 275517438
Omy_106747-707	A/G	fwd: CCGTTAAGAAAGGGTGACATCATGT	VIC: CGATACTCACACTGGCCTG	753	GB: HR504859
		rev: AGATCCATGGCCCCAGTCT	FAM: ATACTCACACCGGCCTG		db: 275517439
Omy_107031-704	C/T	fwd: GGCTTTCGGATACTGAGCAACAA	VIC: TGGACATGATTGCATAGAC	798	GB: HR504860
		rev: TGAACTCACTGTTGGTATGGACTAGA	FAM: CTGGACATGATTACATAGAC		db: 275517440
Omy_107074-217	A/G	fwd: CCGGGCTGTCATGTGACT	VIC: CCCTGGTCTTGACCC	397	GB: HR504861
		rev: CTGCTGACAGGCCTGAGA	FAM: CCTGGCCTTGACCC		db: 275517441
Omy_107285-69	C/G	fwd: GCCCTTGTGACAATGCACTGTTATA	VIC: ATACGTTACTTTTGACCTTGT	704	GB: HR504862
		rev: AGGTCTAGACAGTGTGCCATTTG	FAM: ACGTTACTTTTCACCTTGT		db: 275517442

Table 1.2 Continued

Assay name	Assay targets	Primers (5'-3')	Probes (5'-3')	Cons. length	GeneBank No. dbSNP No.
Omy_107336-170	C/G	fwd: GCCCTCTCACTCATGACATCAAC	VIC: CACTCCTGGGTGCAGAA	471	GB: HR504863
		rev: GCTCCAGCCACTCGCA	FAM: ACTCCTGCGTGCAGAA		db: 275517443
Omy_107607-137	T/G	fwd: TGAGACAACCCAAAGCTTTAAGGAA	VIC: ATGTTCCGACAATAAAT	517	GB: HR504864
		rev: CAACGCACACTATCAGATCACATC	FAM: TGTTCCGACCATAAAT		db: 275517444
Omy_107786-314	G/A	fwd: TGGTTGTCCAAAGCTTTCTTCAGAA	VIC: CACCTCACCCTCCTCC	635	GB: HR504865
		rev: GCTGATACTACAGCATCCAAGGT	FAM: ACCTCACCTTCCTCC		db: 275517445
Omy_107786-584	A/C	fwd: AATGGAATAGTACCCACTTGTGTCC	VIC: CAATGGTAAGATTTTG	635	GB: HR504865
		rev: AGTCAGTCAAGCTCTCTGGAGATAG	FAM: CAATGGTACGATTTTG		db: n/a
Omy_107806-34	C/T	fwd: TCTTTGTCCATGCACATTGATATT	VIC: ATTGGATGTCAGTGTCATT	983	GB: HR504866
		rev: AGCACATTTAGTTAGCAGTGATGGA	FAM: ATTGGATGTCAATGTCATT		db: 275517446
Omy_108007-193	A/G	fwd: GTGAATACCACCCAGGCTTGT	VIC: ATGTTTTCTCCCTACTTAAC	441	GB: HR504867
		rev: GTCCCTTCCCCAGTTTCACTTAATT	FAM: TTTTCTCCCCACTTAAC		db: 275517447
Omy_108735-311	C/T	fwd: GTTTAATCCTGACTTTTCACTTTTGTCATCT	VIC: AACGCCTCGTGACAAT	428	GB: HR504868
		rev: GCGTGCCCTCAATTCCATT	FAM: AACGCCTCATGACAAT		db: 275517448
Omy_108820-85	T/G	fwd: CACCAACAACGTGTAGATTTCCTTAAAATATT	VIC: TTGATATGTGAATTTTG	397	GB: HR504869
		rev: TTTGGTTGGTTGTTTTTATCATTGATACAGTT	FAM: TTGATATGTGCATTTTG		db: 275517449
Omy_109243-222	A/C	fwd: ATGTGCACCTCTTAAATTGTAAGTAAAATGT	VIC: TGTTCATTAAATTGACTTTTT	521	GB: HR504870
		rev: ACCCTATATTCAGTGGCAAGATTGC	FAM: TTCATTAAATGGACTTTTT		db: 275517450
Omy_109390-341	C/T	fwd: ATTACAAACACAAGTCCTCATACAAGTGA	VIC: CATTTTGGCGGTCCAGAA	426	GB: HR504871
		rev: TGTAGGCAACGTTGGTTTATGGT	FAM: CATTTTGGCGATCCAGAA		db: 275517451
Omy_109525-403	A/G	fwd: CCTCATTCTCATTGGTGAGTTGTCT	VIC: CCTACACCTCTTTTTTCCACA	1045	GB: HR504872
		rev: TGTAAGATCTGACCACATGAGTATAACCA	FAM: CCTACACCTCTTTTCTCCACA		db: 275517452
Omy_109651-445	C/T	fwd: CCTGATTTTGCCCACATTTCAAGAA	VIC: CATATGTTAACGTGGGCTAT	615	GB: HR504873
		rev: GCTGTTGTCATATCATCCCGTTAAC	FAM: CATATGTTAACATGGGCTAT		db: 275517453
Omy_109693-461	T/A	fwd: GCCTCACCTGATGCCCATT	VIC: ACGACAGCCACACACAG	474	GB: HR504874
		rev: TGGAGGATTCAGCATTTGGATACC	FAM: ACGACAGCCTCACACAG		db: 275517454
Omy_109874-148	A/G	fwd: GTATGTGTGAGTATGTAATGACTGTATTTAGGA	VIC: ACAGCATTGATTTTGTCACC	392	GB: HR504875
		rev: CTCCTCCCTCAGTGCATTACATTTT	FAM: CAGCATTGATTTCGTCACC		db: 275517455
Omy_109894-185	T/C	fwd: CGGTGTCATTATGGTTGTCATTGTG	VIC: CTCCCTGATCCCCC	581	GB: HR504876
		rev: GGGAGGAATTGGAATGACAGATTAAC	FAM: CTCCCTGGTCCCCC		db: 275517456
Omy_109944-74	T/G	fwd: CCGGGACCAATTGAGAAATCGATAA	VIC: ACGTGACTGTATAGAGACT	116	GB: HR504877
		rev: GGGTTCAAGAGTACACGCCAA	FAM: ACGTGACTGTATCGAGACT		db: 275517457
Omy_110064-419	T/G	fwd: GTGCAAGGGACCTAGCTAATCC	VIC: ACGTTAGCTTTTAATTTC	798	GB: HR504878
		rev: TCTGAACTGACACTGAAGAACAAAGAA	FAM: AACGTTAGCTTTTCATTTC		db: 275517458
Omy_110078-294	A/G	fwd: GCAGTAAATCAGCAGAGACCTACA	VIC: TGTCTACGGATGACTTC	478	GB: HR504879
		rev: CCTTAAGCTCAGATTTAAACGATCAAAACA	FAM: TCTACGGACGACTTC		db: 275517459
Omy_110201-359	T/G	fwd: GGTAAGGCCTGTCTGACTATTTTGA	VIC: TTTGGCTATTGAAATTATACATT	588	GB: HR504880
		rev: AGAGGTCAATGGATGCCAGTTT	FAM: TTGGCTATTGAAATTCTACATT		db: 275517460
Omy_110362-585	G/A	fwd: GCAGCCAAGATGAACGAAAACTTC	VIC: CACCGCCCTGCCCGT	653	GB: HR504881
		rev: CCGGCCTGGGTCTCAATG	FAM: CACCGCCTTGCCCGT		db: 275517461
Omy_110571-386	C/T	fwd: CACTTGGCTCTGCACTAGCA	VIC: CTGTGTAAAATCCATGTCAACA	479	GB: HR504882
		rev: GGGTTGTTAAGAGTCCATTAGAAAGAAC	FAM: TGTGTAAAATCCATATCAACA		db: 275517462
Omy_110689-148	A/C	fwd: GTGTGTGGCAGAGAACTAACTGAT	VIC: CAAATGAACACATTATTTATC	379	GB: HR504884
		rev: GGTTAAGACATTAACATAACACTGGACTCT	FAM: ATGAACACATGATTTATC		db: 275517464
Omy_111005-159	C/T	fwd: ATCTGTCAGACAGTTGTGGATAATGTC	VIC: AGTCAAAAGGGCACAAAA	463	GB: HR504885
		rev: TCGATGACCAACATTGTAGTGTTAAATACA	FAM: AAGTCAAAAGAGCACAAAA		db: 275517465
Omy_111084-526	A/C	fwd: CACCACACCAAGCAACTATTTCATT	VIC: CCAGTGAAATTTATTTTT	709	GB: HR504886
		rev: ACCCAACTACTGTCCCATTTTTCAT	FAM: CAGTGAAATGTATTTTT		db: 275517466
Omy_111383-51	C/T	fwd: CACGCGCAATCTCTCGTTTTAC	VIC: ACCTAGTGCGCTTGCT	495	GB: HR504887
		rev: TCTTTAGGCAACAAGCGTGTCA	FAM: ACCTAGTGCACTTGCT		db: 275517467
Omy_111666-301	T/A	fwd: GGGTGAAAAGAGTGGGACATTTACA	VIC: AGTATAACACAGTAAGACAAT	639	GB: HR504888
		rev: GTCAATTTCAAGGCACCAGACAAT	FAM: AGTATAACACAGTTAGACAAT		db: 275517468
Omy_111681-432	C/T	fwd: GGCGGTTTAAGCAGCAGAAATAC	VIC: TCCCTCTCGGGTGCTG	693	GB: HR504889
		rev: GTGGATCATGCTCGCTAGGT	FAM: CCCTCTCAGGTGCTG		db: 275517469
Omy_112208-328	T/C	fwd: GTCAACAGTTGGACGTAGATGCT	VIC: CTGACAGTGATTATTTTGT	904	GB: HR504890

Table 1.2 Continued

Assay name	Assay targets	Primers (5'-3')	Probes (5'-3')	Cons. length	GeneBank No. dbSNP No.
		rev: CCTTCAGCTTGATCACCTCATAGG	FAM: TGACAGTGATTGTTTTGT		db: 275517470
Omy_112301-202	T/G	fwd: GTAAACCCTGCCCACATAATTAGGT	VIC: AATGCGAAGACAAACT	1146	GB: HR504891
		rev: CTGAGACACTGCTCCAAGGT	FAM: AATGCGAAGCCAAACT		db: 275517471
Omy_112820-82	G/A	fwd: CCTTTCCTTTTGCATTTCCTCTACTTATTTATTT	VIC: CGCCGCCAAGTTA	393	GB: HR504892
		rev: AAATGAACTCACGTTGACCTCTGA	FAM: CGCCGCTAAGTTA		db: 275517472
Omy_112876-45	T/C	fwd: GGACTACATGAAGGCGTGAGT	VIC: TTTTAGTGACGAGTGTCTG	805	GB: HR504893
		rev: ATCAGTCCTAGCCCAAACACATG	FAM: TAGTGACGGGTGTCTG		db: 275517473
Omy_113109-205	T/G	fwd: GTGGGCACTGTTACACAAAGTTC	VIC: CGTCATCTTAAATTATCTTTG	416	GB: HR504894
		rev: CCAGTCAACTTACAAACAAGCCATT	FAM: CGTCATCTTAAATTCTCTTTG		db: 275517474
Omy_113128-73	C/G	fwd: CCTCCTACTCTGATCTAAAGATTACAGAA	VIC: TGGCAGGGTTTCCGG	374	GB: HR504895
		rev: TTCTCTGCCCTCTCGATTTTGG	FAM: TGGCAGGCTTTCCGG		db: 275517475
Omy_113242-163	T/C	fwd: TGGTGGACTGATCTGATGATGAAAG	VIC: TCTGAGACAACACGCTAT	389	GB: HR504896
		rev: CCTCGTCCATATTTTCCTCCTCAA	FAM: CTGAGACAACGCGCTAT		db: 275517476
Omy_113490-159	C/T	fwd: CATAGTACATTTACAGATAATGTTTTAAAGTGCATGT	VIC: CATCTGTTTTGGTTTAGC	288	GB: HR504897
		rev: CGAGATACCAAAATGCCACAGTTACAT	FAM: CATCTGTTTTAGTTTAGC		db: 275517477
Omy_114315-438	T/G	fwd: CCTCACCGATCTAGTCAACTTCATC	VIC: TTATGGGCTTAAGGGTC	555	GB: HR504898
		rev: AGGAGGCTGAGGGAGATTCTAG	FAM: TTATGGGCTTACGGGTC		db: 275517478
Omy_114448-87	C/T	fwd: GCCGAAAGGTAAAATCCACAAATCC	VIC: TGGTTGATCGAACATTT	530	GB: HR504899
		rev: GGACTAGGCTAACAGGAGAAGCT	FAM: TGGTTGATCAAACATTT		db: 275517479
Omy_114587-480	T/G	fwd: CAGATTACGTTATTACGTTTGGGAAATTTTTAAGT	VIC: CCTGTCCAAAATTGT	1266	GB: HR504900
		rev: GTGAAAGAGTGGGAAATATAATTATAAGGTCAGA	FAM: CCTGTCCACAATTGT		db: 275517480
Omy_114976-223	T/G	fwd: GACAAACAGCACTTCATTGCAGTAA	VIC: ACCGATGGAACAATC	735	GB: HR504901
		rev: GTTGCTCCAGCACCAGGT	FAM: CCGATGGCACAATC		db: 275517481
Omy_115987-812	C/T	fwd: GAGCTCCTGAAGACCTATAAGAATGTT	VIC: CTGAAAAGACTGCTCCAC	1166	GB: HR504902
		rev: GGTCGAGGAAGAGCTCAATGC	FAM: CTGAAAAAACTGCTCCAC		db: 275517482
Omy_116104-229	T/C	fwd: GCTAGAAGATAACAGGCCACACT	VIC: TGACAAGTTTAAGCTTG	513	GB: HR504903
		rev: ATGGTATTCAATGGCATTTTCAGTTTCAAA	FAM: TGACAAGTTTAGGCTTG		db: 275517483
Omy_116362-467	T/G	fwd: CTGGATCCAAGAGGCTGTTCT	VIC: CTCACCTGAATCCAG	508	GB: HR504904
		rev: TGCCTGCTATAGTTCCATGTCAAAA	FAM: CTCACCTGCATCCAG		db: 275517484
Omy_116733-349	C/T	fwd: GAAATGGACATGCCTACAAATTGCT	VIC: AGAGAATCTGATAGTATTTC	641	GB: HR504905
		rev: GATGTGATCAGTTTAGGCAAGGC	FAM: AGAGAATCTGATAATATTTC		db: 275517485
Omy_116938-264	A/G	fwd: GTTCATTCATGTTGAAGTGCGACAT	VIC: CCTTGTCTCAATTTTTCCTCT	530	GB: HR504906
		rev: CTCTGCATGCTCCCATCCT	FAM: CTTGTCTCAATTTCTCCTCT		db: 275517486
Omy_117242-419	G/A	fwd: GTCTTCTCTCTTTCTCTCCCTCTCT	VIC: CCTCCCTGCCTCCCT	479	GB: HR504907
		rev: CCACTGGCCTTCAATTGTAACAG	FAM: CCTCCCTGTCTCCCT		db: 275517487
Omy_117259-96	T/C	fwd: CAAGGGAAGAGCTCTGAGATGAG	VIC: CGTCATGCCATCATGT	409	GB: HR504908
		rev: GGGATCAGTGGCAGGTAGAG	FAM: CGTCATGCCGTCATGT		db: 275517488
Omy_117286-374	A/T	fwd: TGATGTGTTGTTCCTCATGGCTTA	VIC: CTTTCCTCATCATACTCTATGG	453	GB: HR504909
		rev: CTGTGCATTTATTCTTGTGATGCTAGG	FAM: TCCTCATCATACACTATGG		db: 275517489
Omy_117370-400	A/G	fwd: TGCAAACACAGAGGAAAGGGATTT	VIC: CAACTCCAATGAATTAA	596	GB: HR504910
		rev: GGCTTATTTGTTCCGTACTTGCATT	FAM: AACTCCAACGAATTAA		db: 275517490
Omy_117432-190	C/T	fwd: GGAGAACGCCTTGAGGTTGT	VIC: TCATGGTGGATCCTGG	441	GB: HR504911
		rev: TGCCTCATCCTTGGGACTGAT	FAM: TCATGGTGAATCCTGG		db: 275517491
Omy_117540-259	T/G	fwd: GGCAGGTTAACACAGTCATCTACTATAAA	VIC: TGTCACTTCAAAGTTTG	575	GB: HR504912
		rev: CAGCATGTTGCTTTAATCCTTCACA	FAM: TGTCACTTCAACGTTTG		db: 275517492
Omy_117549-316	A/G	fwd: CCAGTACCCTTACATCTGAGAACCA	VIC: CTGCCCTTGCTGGC	425	GB: HR504913
		rev: GGCCTTGGTTGTAGTTGTCACT	FAM: TGCCCCTGCTGGC		db: 275517493
Omy_117743-127	C/T	fwd: ACCTGCACCTTGTAAATAATTTATATAGTAGCTAAATAATT	VIC: ACATACAGAACGTTCACTG	477	GB: HR504914
		rev: GCCTGCCTGTGAACAACAC	FAM: ACATACAGAACATTCACTG		db: 275517494
Omy_117815-81	C/T	fwd: CTGCTTTATGCACACCACATTGT	VIC: CTATACGGAGACCAGC	402	GB: HR504915
		rev: GCTCTTTCTGGAGAACAAGGTACTG	FAM: CTATACGGAAACCAGC		db: 275517495
Omy_118175-396	T/A	fwd: AGGCTTCACACACACATGCA	VIC: CTCTTGCAGACATACCCGTA	463	GB: HR504916
		rev: GACGCGCAACCTCTAGATTATACTT	FAM: CTCTTGCAGACATTCCCGTA		db: 275517496
Omy_118205-116	A/G	fwd: CTGCGGTGGGCTACACA	VIC: CTACTGAGGCTGAGTGCT	485	GB: HR504917
		rev: CGCAGCTGCGGATGAG	FAM: TACTGAGGCCGAGTGCT		db: 275517497

Table 1.2 Continued

Assay name	Assay targets	Primers (5'-3')	Probes (5'-3')	Cons. length	GeneBank No. dbSNP No.
Omy_118654-91	A/G	fwd: CAGCGTAGACCGTTTCCTCATTAT	VIC: TCAGCTTGTCTTGCCGC	454	GB: HR504918
		rev: GCGCCGATGAGCAGCTT	FAM: CAGCTTGTCCTGCCGC		db: 275517498
Omy_118938-341	A/T	fwd: GAGGGACAGACTTCAAGATTTCATGA	VIC: TGTTGTTCAGATTGTAAAAA	625	GB: HR504919
		rev: AGTCATCATAAAGACTTGTTCATTAAGGAAGG	FAM: TGTTGTTCAGATAGTAAAAA		db: 275517499
Omy_119108-357	T/C	fwd: GGTAGAAGCAGCCCATGCA	VIC: CGCGTCCAAGCAG	949	GB: HR504920
		rev: TGTGGCAAGGACATGTGTGA	FAM: CGCGTCCAGGCAG		db: 275517500
Omy_119892-365	T/G	fwd: GGTTATAGGTTCGTCACCATCCAAA	VIC: AATTCTACCTACAGCTAACA	755	GB: HR504921
		rev: TTGTCTGTGGTGTTATGTCTAATTTCAAG	FAM: ATTCTACCTACCGCTAACA		db: 275517501
Omy_120255-332	A/T	fwd: GGCTACAGGGACTTTACAATGGG	VIC: ACTATGCCATGAAGTTA	601	GB: HR504922
		rev: GCTAGCTAACATTGAAGGGTGGAAT	FAM: ACTATGCCAAGAAGTTA		db: 275517502
Omy_120950-569	T/G	fwd: TCACACTCAGATTATTGTGGCGATT	VIC: ATTGTTTAACCTAAAAGCTT	759	GB: HR504923
		rev: GCTGACTCATAAAAATGTTGGTAATGCT	FAM: TGTTTAACCTACAAGCTT		db: 275517503
Omy_121006-131	T/G	fwd: ACAGTGAATCAGCGGAGAAACA	VIC: TTCGTACGAGACCAAAG	505	GB: HR504924
		rev: AGTCCGTTTCCTGTTAGTGTAAGC	FAM: TCGTACGAGCCCAAAG		db: 275517504
Omy_121713-115	T/A	fwd: TGTGACAGAGCCAAGGAAAACC	VIC: TCAGGTTGAGTATTGC	501	GB: HR504925
		rev: TGGGCTAGTGAGGGAGTGA	FAM: TCAGGTTGTGTATTGC		db: 275517505
Omy_123044-128	C/T	fwd: CTGGGTGAGTGAGTTGACTATACAC	VIC: ATTTCTGGCGGTCCGG	784	GB: HR504926
		rev: CGGGTGTGCATGAGAAAATGAC	FAM: ATTTCTGGCAGTCCGG		db: 275517506
Omy_123048-119	C/T	fwd: ATGTATCTGGTGCATTGGGATGATT	VIC: ACTTGCCCGATACTT	797	GB: HR504927
		rev: ACAGCCACATGTACAGGGAAAAA	FAM: ACTTGCCCAATACTT		db: 275517507
Omy_123921-144	T/C	fwd: AACTCTGAAGTGGGATGTGATGTTC	VIC: CTAAGGTTCAGGACTTGGA	1045	GB: HR504928
		rev: GGATGATGTTACAAAAGGAGAGCATGT	FAM: AAGGTTCAGGGCTTGGA		db: 275517508
Omy_124774-530	A/T	fwd: AGTACCACCGCCGTCTGATATAT	VIC: CAAATAAAAGGCTAAATAAA	705	GB: HR504929
		rev: CCAGAGCAAAGCATGTCCTCAAATA	FAM: AAATAAAAGGCAAAATAAA		db: 275517509
Omy_125998-61	T/G	fwd: GGTGTCCAGCCACAGTACAG	VIC: TGACCTCCATCCCCC	459	GB: HR504930
		rev: TGTTCCTTTATTGGGCCTGCATA	FAM: ATGACCTCCCTCCCCC		db: 275517510
Omy_126160-242	T/G	fwd: CAAGGGAGTGACCGGAATGTTATAT	VIC: CAATCATGTGTTAACACTAA	648	GB: HR504931
		rev: GCCCAGACATTTACAGCAGTATCA	FAM: ATCATGTGTTCACACTAA		db: 275517511
Omy_127236-583	C/G	fwd: TGGATCAAGACAGATTTCCCCTACA	VIC: ATTGTGAAACGGCCCCT	685	GB: HR504932
		rev: GCCACCAGTGAGATGTCTTTGAAA	FAM: ATTGTGAAACCGCCCCT		db: 275517512
Omy_127510-920	C/T	fwd: GTGTTATGCCAACAAGGCTTGT	VIC: AACAAATAACAGACGACATTA	1182	GB: HR504933
		rev: TTTGACAATATCAATATCATGAAAATGTTTGTGAGT	FAM: ACAAATAACAGACAACATTA		db: 275517513
Omy_127645-308	A/T	fwd: ACACTGATATTAACATGGCACAAGTCA	VIC: AAGTTTGTTACATATTTTG	401	GB: HR504934
		rev: CAGGGCCGGTCGTAGATTTT	FAM: TTTGTTACAAATTTTG		db: 275517514
Omy_127760-385	A/T	fwd: CGGCTATTCTCGCGTAAAAGCT	VIC: TCCTTATCCAAAATTATTGTGC	756	GB: HR504935
		rev: AAATGCAACCAGAAACGGAATGTC	FAM: CTTATCCAAAATAATTGTGC		db: 275517515
Omy_128302-430	C/T	fwd: GTATGGCATTTTTGTTCCCAAGGT	VIC: CATCATCGTAAATCAG	1025	GB: HR504936
		rev: CATGTGGTTGCCCTCCTTATAGAG	FAM: CATCATCATAAATCAG		db: 275517516
Omy_128693-755	A/C	fwd: GATACACTCTACTGACTAGTCCATCCA	VIC: CTCTGACCATTATTTTGTC	869	GB: HR504937
		rev: GTCCTGAAAGAGAGAAACACAGACA	FAM: CTGACCATTAGTTTGTC		db: 275517517
Omy_128851-273	T/A	fwd: GTACAGATGAATGTGTTTTATTTGGCATTG	VIC: CCTGTCAAATAAAG	348	GB: HR504938
		rev: CTGCCCATCAAGGTCTTCATCTTAT	FAM: CCCTGTCTAATAAAG		db: 275517518
Omy_128923-433	T/C	fwd: ACGTTTCTTTGGGCTGAGACTTATT	VIC: CTTCATTTTCATTCACTGTTTT	505	GB: HR504939
		rev: CTATGTCCTTGGCAGAAGTCTACA	FAM: CATTTTCATTCGCTGTTTT		db: 275517519
Omy_128996-481	T/G	fwd: CTCATCCACACTGTACAGTACAAGT	VIC: CTTGTGGTTGAGGTTTG	515	GB: HR504940
		rev: CATGCCTTCGTCTCATCAATAACAC	FAM: TTGTGGTTGCGGTTTG		db: 275517520
Omy_129170-794	T/G	fwd: GTTAGAAACCATGACTCACCATCCA	VIC: CCCTGTGGAGTGTCAG	830	GB: HR504941
		rev: CTGTAGCAGTGATGCTATGGAATAGG	FAM: CCTGTGGCGTGTCAG		db: 275517521
Omy_129870-756	C/T	fwd: TCGTTATTTTGCCTCGCGGTA	VIC: ACAGGTATTTCGTGAAATG	965	GB: HR504942
		rev: TCCCATGAAGATGTATACATGTTTTGTGA	FAM: CAGGTATTTCATGAAATG		db: 275517522
Omy_130295-98	A/C	fwd: GGGACCACAGAATATTTTTCTTGTTCAT	VIC: CTTATGCCTTTTCTAATTCTGTA	583	GB: HR504943
		rev: TGGACAGAATGTTCTACAAGTTGCA	FAM: TTATGCCTTTTCTAAGTCTGTA		db: 275517523
Omy_130524-160	C/G	fwd: CGAAGGTAGCGATTGGTCGTT	VIC: ATGGCTTGATCCTCA	388	GB: HR504944
		rev: TGTCTGTTCTGCTGTGTGCTT	FAM: ATGGCTTCATCCTCA		db: 275517524
Omy_130720-100	C/T	fwd: CGGTCATTGTAAATGTCAACGGTTT	VIC: ACCTGTCCCGTTCCCA	547	GB: HR504945

Table 1.2 Continued

Assay name	Assay targets	Primers (5'-3')	Probes (5'-3')	Cons. length	GeneBank No. dbSNP No.
		rev: TGCTTGCATGTTCTTGGTGTAGTA	FAM: CCTGTCCCATTCCCA		db: 275517525
Omy_131460-646	C/T	fwd: GTGAAAAGGAATGGAGGAGTACAGT	VIC: AATAAAGCAGAATTTGTTACTG	1276	GB: HR504946
		rev: TGCTAGGACAGGAAGATCATTTGTG	FAM: AAAGCAGAATTTATTACTG		db: 275517526
Omy_131965-120	C/T	fwd: AGAGATACATTAAAGCTGTGCTCATTCA	VIC: CATTGTAAACGACCATTTT	240	GB: HR504947
		rev: GCAGAGTTGCTTCAAAACTGTTAGT	FAM: CATTGTAAACAACCATTTT		db: 275517527

Mean minor allele frequency (MAF) averaged 0.197 over all loci, with a high of 0.258 in the Sacramento River-Battle Creek and the Russian River populations and a low of 0.112 in the McCloud River-Butcherknife Creek population. The proportion of polymorphic loci averaged 81.5% and varied from 97.1% in Battle Creek to 56.8% in Butcherknife Creek (Table 1.3). Expected and observed heterozygosity were generally very similar within each test sample, never differing more than 0.026 (*i.e.* 2.6%). Observed heterozygosity varied between 0.342 in Battle Creek and 0.0001 in Butcherknife Creek (Table 1.4). Thus, all measures of genetic variability were consistent in identifying the Sacramento River-Battle Creek population as the most diverse and the McCloud River-Butcherknife Creek population as the least diverse. Mean F_{ST} was 0.206 and ranged from 0 to 0.629 at different loci (Table 1.4).

Table 1.3: Allele frequency of 139 SNP assays in 11 natural-origin steelhead populations, four *O. mykiss* hatchery strains, and two introduced *O. mykiss* populations in the southern hemisphere. Allele frequency (AF) reported for all groups is the minor allele ($p < 0.5$) in Willamette River-Wiley Creek (MAF). A_RH: American River Hatchery; BCH: Big Creek Hatchery; HCH: Hot Creek Hatchery; FH: Fillmore Hatchery.

	Willamette River, Wiley Creek	Klamath River, Buckboard Creek	Klamath River, Kelsey Creek	Goose Lake, Bauers Creek	McCloud River, Butcherknife Creek	Sacramento River, Battle Creek	Eel River, Middle Fork (summer)	Russian River	Scott Creek	San Lorenzo River	Carmel River	Eagle Lake Strain, A_RH	San Lorenzo Strain, BCH	Kamloops Strain, HCH	Mount Whitney Strain, FH	Lake Taupo, New Zealand	Río Santa Cruz, Argentina
N	23	15	23	16	16	24	24	47	46	15	24	16	9	15	16	23	24
Assay Name	MAF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF
OMGH1PROM1-SNP1	0.04	0.85	0.20	0.72	0.06	0.50	0.31	0.31	0.26	0.10	0.23	0.09	0.11	0.07	0.34	0.73	0.30
Omy_95318-147	0.41	1.00	0.91	0.44	1.00	0.85	0.94	0.61	0.50	0.43	0.46	0.87	0.56	0.20	0.53	0.35	0.76
Omy_95442-108	0.00	0.70	0.04	0.09	0.00	0.15	0.02	0.29	0.17	0.18	0.37	0.03	0.13	0.00	0.00	0.00	0.02
Omy_95489-423	0.00	0.27	0.02	0.06	0.81	0.39	0.56	0.63	0.66	0.63	0.56	0.09	0.63	0.13	0.88	0.55	0.65
Omy_96158-277	0.00	0.23	0.18	0.06	0.72	0.35	0.02	0.28	0.46	0.27	0.27	0.34	0.28	0.13	0.25	0.24	0.23
Omy_96222-125	0.09	0.10	0.60	0.16	0.22	0.11	0.17	0.16	0.26	0.17	0.38	0.00	0.17	0.07	0.03	0.00	0.25
Omy_96529-231	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.11	0.20	0.07	0.00	0.22	0.00	0.00	0.00	0.00
Omy_96899-148	0.15	0.00	0.09	0.75	0.00	0.46	0.33	0.60	0.61	0.57	0.56	0.63	0.50	0.43	0.69	0.37	0.98
Omy_97077-73	0.13	0.00	0.04	0.03	0.00	0.28	0.50	0.18	0.25	0.23	0.39	0.00	0.17	0.00	0.31	0.11	0.00
Omy_97660-230	0.36	0.00	0.00	0.22	0.00	0.17	0.27	0.21	0.06	0.03	0.11	0.09	0.06	0.70	0.19	0.30	0.04
Omy_97865-196	0.00	0.00	0.00	0.06	0.00	0.09	0.00	0.03	0.06	0.00	0.02	0.38	0.00	0.00	0.27	0.00	0.00
Omy_97954-618	0.00	0.94	0.02	0.78	0.50	0.72	0.50	0.53	0.36	0.33	0.65	0.16	0.28	0.37	0.72	0.76	0.25
Omy_98188-405	0.22	0.03	0.05	0.06	0.03	0.20	0.06	0.22	0.38	0.21	0.44	0.19	0.33	0.30	0.06	0.18	0.00
Omy_98409-549	0.09	0.90	0.26	0.00	0.03	0.31	0.33	0.38	0.49	0.43	0.31	0.25	0.44	0.00	0.00	0.06	0.15
Omy_98683-165	0.04	0.40	0.00	0.09	0.00	0.39	0.52	0.56	0.66	0.60	0.25	0.25	0.56	0.13	0.41	0.37	0.94
Omy_99300-202	0.17	0.00	0.02	0.00	0.00	0.22	0.00	0.18	0.17	0.03	0.15	0.16	0.00	0.50	0.31	0.00	0.02
Omy_100771-63	0.11	0.13	0.00	0.00	0.03	0.57	0.83	0.46	0.25	0.43	0.27	0.63	0.38	0.07	0.00	0.87	0.19
Omy_100974-386	0.02	0.60	0.19	0.63	0.16	0.39	0.19	0.27	0.17	0.07	0.21	0.22	0.06	0.10	0.47	0.30	0.17
Omy_101119-554	0.00	0.13	0.07	0.03	0.69	0.33	0.00	0.00	0.00	0.03	0.00	0.28	0.06	0.07	0.06	0.57	0.33
Omy_101341-188	0.00	0.00	0.00	0.00	0.00	0.04	0.27	0.07	0.24	0.10	0.08	0.00	0.06	0.00	0.00	0.15	0.13
Omy_101554-306	0.02	0.13	0.05	0.13	0.00	0.22	0.35	0.52	0.48	0.67	0.60	0.69	0.61	0.27	0.56	0.22	0.04
Omy_101704-329	0.00	0.00	0.00	0.00	0.00	0.30	0.00	0.09	0.00	0.00	0.00	0.06	0.00	0.13	0.47	0.48	0.04
Omy_101770-410	0.13	0.93	0.18	0.00	0.16	0.22	0.21	0.31	0.21	0.23	0.33	0.38	0.31	0.23	0.13	0.07	0.10
Omy_101832-195	0.20	0.80	0.87	0.72	0.97	0.72	0.71	0.49	0.43	0.47	0.42	0.91	0.25	0.57	0.97	0.22	0.63
Omy_101993-189	0.02	1.00	0.98	1.00	0.28	0.42	0.96	0.65	0.48	0.53	0.63	0.41	0.31	0.70	0.13	0.04	0.46
Omy_102213-204	0.02	0.20	0.97	0.88	0.13	0.30	0.13	0.12	0.13	0.11	0.06	0.77	0.00	0.03	0.31	0.68	0.69

Table 1.3 Continued

		Willamette River, Wiley Creek	Klamath River, Buckboard Creek	Klamath River, Kelsey Creek	Goose Lake, Bauers Creek	McCloud River, Butcherknife Creek	Sacramento River, Battle Creek	Eel River, Middle Fork (summer)	Russian River	Scott Creek	San Lorenzo River	Carmel River	Eagle Lake Strain, A_RH	San Lorenzo Strain, BCH	Kamloops Strain, HCH	Mount Whitney Strain, FH	Lake Taupo, New Zealand	Río Santa Cruz, Argentina
	N	23	15	23	16	16	24	24	47	46	15	24	16	9	15	16	23	24
Assay Name		MAF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF
Omy_102420-634		0.15	0.23	0.30	0.00	0.63	0.24	0.56	0.59	0.66	0.77	0.69	0.06	0.67	0.47	0.28	0.61	0.67
Omy_102457-423		0.09	0.80	0.58	0.84	1.00	0.28	0.36	0.51	0.79	0.50	0.58	0.56	0.56	0.23	0.42	0.44	0.77
Omy_102505-102		0.25	0.03	0.13	0.00	0.03	0.02	0.23	0.14	0.23	0.13	0.15	0.00	0.22	0.07	0.16	0.11	0.02
Omy_102510-682		0.04	0.83	0.13	0.00	0.00	0.22	0.13	0.31	0.07	0.13	0.25	0.06	0.11	0.00	0.31	0.28	0.00
Omy_102867-443		0.00	0.07	0.00	0.06	0.00	0.22	0.29	0.46	0.73	0.70	0.58	0.00	0.72	0.03	0.78	0.28	0.88
Omy_103350-395		0.09	0.67	0.18	0.06	0.00	0.65	0.63	0.42	0.28	0.47	0.48	0.77	0.44	0.17	0.41	0.50	0.54
Omy_103577-379		0.18	0.20	0.05	0.00	0.00	0.07	0.08	0.46	0.29	0.40	0.33	0.03	0.39	0.00	0.31	0.22	0.10
Omy_103705-558		0.24	0.03	0.15	0.28	0.78	0.13	0.25	0.38	0.34	0.27	0.44	0.09	0.44	0.20	0.00	0.44	0.10
Omy_103713-53		0.00	0.37	0.02	0.00	0.09	0.54	0.23	0.44	0.46	0.47	0.23	0.16	0.61	0.18	0.25	0.67	0.75
Omy_104519-624		0.28	0.90	0.26	0.06	0.03	0.33	0.19	0.53	0.73	0.73	0.77	0.38	0.61	0.47	0.38	0.41	0.52
Omy_104569-114		0.30	0.20	0.52	0.00	0.00	0.28	0.29	0.29	0.15	0.17	0.28	0.00	0.28	0.04	0.13	0.13	0.27
Omy_105075-162		0.26	0.00	0.41	0.88	0.47	0.15	0.10	0.10	0.32	0.33	0.44	0.00	0.39	0.03	0.00	0.00	0.23
Omy_105105-448		0.35	0.10	0.30	0.81	0.84	0.76	0.60	0.84	0.92	1.00	0.79	0.94	1.00	0.43	0.53	1.00	0.65
Omy_105115-367		0.02	0.87	0.21	0.00	0.03	0.13	0.23	0.26	0.29	0.10	0.08	0.16	0.06	0.13	0.07	0.15	0.38
Omy_105235-713		0.00	0.00	0.00	0.00	0.00	0.07	0.21	0.15	0.24	0.10	0.13	0.00	0.06	0.00	0.00	0.25	0.17
Omy_105385-406		0.39	0.43	0.61	0.59	0.34	0.38	0.30	0.32	0.21	0.27	0.40	0.38	0.25	0.70	0.75	0.03	0.29
Omy_105386-347		0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.33	0.17	0.38	0.66	0.11	0.00	0.00	0.09	0.00
Omy_105401-363		0.00	0.00	0.02	0.00	0.78	0.17	0.35	0.39	0.44	0.50	0.58	0.00	0.50	0.00	0.00	0.21	0.15
Omy_105407-74		0.00	0.20	0.43	0.41	0.03	0.39	0.42	0.73	0.64	0.70	0.73	0.63	0.56	0.00	0.50	0.33	0.54
Omy_105714-265		0.05	0.13	0.15	0.44	0.41	0.61	0.75	0.84	0.70	0.63	0.77	0.19	0.50	0.83	0.75	0.74	0.63
Omy_105897-101		0.00	0.17	0.02	0.66	0.00	0.21	0.00	0.01	0.01	0.00	0.02	0.09	0.00	0.10	0.09	0.17	0.08
Omy_106172-332		0.00	0.33	0.35	0.00	0.38	0.13	0.08	0.17	0.00	0.03	0.00	0.00	0.00	0.03	0.13	0.09	0.13
Omy_106313-445		0.00	0.00	0.35	0.00	0.63	0.63	0.17	0.50	0.61	0.50	0.79	0.03	0.39	0.13	0.59	0.25	0.52
Omy_106560-58		0.00	0.07	0.02	0.88	0.47	0.35	0.19	0.20	0.22	0.03	0.20	0.59	0.00	0.03	0.22	0.55	0.17
Omy_106747-707		0.48	0.73	0.52	0.84	0.56	0.39	0.54	0.39	0.32	0.46	0.38	0.44	0.44	0.77	0.56	0.46	0.61
Omy_107031-704		0.41	0.73	0.11	0.72	0.31	0.38	0.08	0.03	0.04	0.00	0.00	0.13	0.00	0.87	0.38	0.17	0.48
Omy_107074-217		0.00	0.00	0.02	0.00	0.00	0.15	0.54	0.70	0.89	0.90	0.79	0.06	0.89	0.00	0.66	0.24	0.96
Omy_107285-69		0.35	0.73	0.09	0.81	0.00	0.57	0.21	0.28	0.27	0.33	0.33	0.13	0.33	0.23	0.53	0.54	0.37
Omy_107336-170		0.05	0.07	0.11	0.50	0.00	0.24	0.00	0.01	0.00	0.00	0.00	0.28	0.00	0.27	0.00	0.39	0.08
Omy_107607-137		0.48	0.93	0.98	1.00	0.78	0.87	0.63	0.86	0.81	0.83	0.92	0.97	0.78	0.83	0.84	0.94	1.00

Table 1.3 Continued

		Willamette River, Wiley Creek	Klamath River, Buckboard Creek	Klamath River, Kelsey Creek	Goose Lake, Bauers Creek	McCloud River, Butcherknife Creek	Sacramento River, Battle Creek	Eel River, Middle Fork (summer)	Russian River	Scott Creek	San Lorenzo River	Carmel River	Eagle Lake Strain, A_RH	San Lorenzo Strain, BCH	Kamloops Strain, HCH	Mount Whitney Strain, FH	Lake Taupo, New Zealand	Río Santa Cruz, Argentina
	N	23	15	23	16	16	24	24	47	46	15	24	16	9	15	16	23	24
Assay Name		MAF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF
Omy_107786-314		0.09	0.97	0.89	0.13	1.00	0.28	0.90	0.71	0.88	0.83	0.79	0.70	0.89	0.10	0.59	0.17	0.60
Omy_107786-584		0.03	0.00	0.12	0.88	0.00	0.12	0.11	0.21	0.09	0.11	0.12	0.31	0.06	0.50	0.25	0.09	0.37
Omy_107806-34		0.40	0.70	0.57	0.97	0.81	0.39	0.02	0.05	0.08	0.20	0.02	0.88	0.22	0.50	0.84	0.76	0.75
Omy_108007-193		0.24	0.93	0.11	0.31	0.44	0.35	0.48	0.11	0.01	0.07	0.04	0.75	0.06	0.23	0.50	0.41	0.31
Omy_108735-311		0.41	0.10	0.61	0.03	0.31	0.52	0.63	0.72	0.80	0.67	0.83	0.25	0.56	0.17	0.59	0.33	0.58
Omy_108820-85		0.02	0.10	0.13	0.03	0.34	0.15	0.00	0.01	0.00	0.00	0.00	0.33	0.00	0.10	0.16	0.22	0.08
Omy_109243-222		0.09	0.23	0.40	0.84	0.03	0.48	0.13	0.13	0.12	0.23	0.08	0.81	0.17	0.07	0.53	0.41	0.52
Omy_109390-341		0.00	0.10	0.00	0.00	0.00	0.26	0.28	0.13	0.20	0.12	0.21	0.19	0.06	0.23	0.47	0.15	0.13
Omy_109525-403		0.48	0.93	0.35	0.94	0.06	0.74	0.27	0.45	0.38	0.21	0.25	0.91	0.28	0.50	0.56	0.80	0.13
Omy_109651-445		0.00	0.00	0.00	0.03	0.00	0.02	0.02	0.18	0.24	0.17	0.29	0.00	0.22	0.00	0.44	0.09	0.13
Omy_109693-461		0.02	0.53	0.09	0.16	0.00	0.09	0.52	0.51	0.60	0.75	0.67	0.38	0.79	0.00	0.44	0.11	0.54
Omy_109874-148		0.07	0.00	0.00	0.00	0.00	0.02	0.08	0.14	0.24	0.40	0.25	0.00	0.44	0.00	0.09	0.11	0.00
Omy_109894-185		0.24	0.62	0.28	0.91	0.09	0.46	0.13	0.20	0.14	0.03	0.21	0.66	0.00	0.43	0.09	0.90	0.71
Omy_109944-74		0.00	0.00	0.07	0.06	0.00	0.41	0.17	0.11	0.03	0.03	0.02	0.84	0.00	0.00	0.16	0.26	0.06
Omy_110064-419		0.00	0.27	0.20	0.44	0.84	0.30	0.09	0.11	0.30	0.53	0.29	0.09	0.50	0.23	0.38	0.28	0.63
Omy_110078-294		0.11	0.83	0.22	0.59	0.00	0.54	0.90	0.75	0.69	0.83	0.85	0.94	0.83	0.20	0.70	0.68	0.73
Omy_110201-359		0.30	0.17	0.52	0.00	0.13	0.20	0.54	0.46	0.50	0.40	0.63	0.19	0.38	0.10	0.34	0.54	0.42
Omy_110362-585		0.11	0.00	0.04	0.19	0.06	0.39	0.44	0.45	0.33	0.47	0.54	0.13	0.56	0.20	0.00	0.15	0.38
Omy_110571-386		0.09	0.00	0.28	0.00	0.06	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.70	0.09	0.00	0.06
Omy_110689-148		0.50	0.20	0.48	0.44	0.03	0.39	0.46	0.61	0.49	0.57	0.54	0.31	0.72	0.17	0.56	0.46	0.33
Omy_111005-159		0.39	0.07	0.88	0.28	1.00	0.44	0.23	0.14	0.01	0.07	0.17	0.41	0.00	0.00	0.72	0.15	0.52
Omy_111084-526		0.13	0.00	0.04	0.00	0.00	0.15	0.02	0.04	0.11	0.10	0.19	0.38	0.11	0.10	0.00	0.00	0.21
Omy_111383-51		0.28	0.97	0.59	0.25	0.75	0.57	0.71	0.85	0.85	0.87	0.94	0.53	0.83	0.37	0.72	0.61	0.65
Omy_111666-301		0.02	0.10	0.28	0.00	0.88	0.54	0.58	0.54	0.46	0.50	0.52	0.50	0.38	0.23	0.44	0.67	0.58
Omy_111681-432		0.00	0.03	0.04	0.88	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.03	0.03	0.37	0.35
Omy_112208-328		0.31	0.33	0.54	0.78	0.16	0.70	0.27	0.38	0.34	0.40	0.44	0.81	0.44	0.20	0.25	0.48	0.57
Omy_112301-202		0.09	0.50	0.64	0.38	0.81	0.74	0.75	0.98	0.95	1.00	1.00	0.31	1.00	0.13	0.53	0.85	0.79
Omy_112820-82		0.04	0.00	0.71	0.91	0.78	0.59	0.85	0.64	0.82	0.83	1.00	0.83	0.78	0.07	0.75	0.50	0.54
Omy_112876-45		0.48	0.43	0.78	0.69	0.28	0.26	0.56	0.48	0.50	0.43	0.40	0.22	0.50	0.40	0.25	0.26	0.41
Omy_113109-205		0.00	0.00	0.00	0.00	0.00	0.33	0.04	0.48	0.30	0.40	0.27	0.25	0.44	0.10	0.56	0.78	0.21

Table 1.3 Continued

		Willamette River, Wiley Creek	Klamath River, Buckboard Creek	Klamath River, Kelsey Creek	Goose Lake, Bauers Creek	McCloud River, Butcherknife Creek	Sacramento River, Battle Creek	Eel River, Middle Fork (summer)	Russian River	Scott Creek	San Lorenzo River	Carmel River	Eagle Lake Strain, A_RH	San Lorenzo Strain, BCH	Kamloops Strain, HCH	Mount Whitney Strain, FH	Lake Taupo, New Zealand	Río Santa Cruz, Argentina
	N	23	15	23	16	16	24	24	47	46	15	24	16	9	15	16	23	24
Assay Name		MAF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF
Omy_113128-73		0.00	0.00	0.15	0.03	0.03	0.07	0.33	0.55	0.46	0.50	0.63	0.09	0.38	0.03	0.06	0.26	0.10
Omy_113242-163		0.00	0.20	0.07	0.81	0.00	0.09	0.00	0.01	0.04	0.00	0.00	0.41	0.00	0.03	0.00	0.11	0.02
Omy_113490-159		0.36	0.70	0.50	0.81	0.97	0.17	0.17	0.01	0.00	0.00	0.00	0.34	0.00	0.37	0.41	0.02	0.42
Omy_114315-438		0.04	0.07	0.15	0.16	0.38	0.35	0.13	0.33	0.41	0.30	0.33	0.06	0.22	0.13	0.16	0.00	0.13
Omy_114448-87		0.00	0.00	0.32	0.00	0.00	0.41	0.69	0.42	0.82	0.90	0.96	0.06	0.83	0.00	0.56	0.44	0.56
Omy_114587-480		0.17	0.10	0.07	0.13	0.00	0.02	0.00	0.19	0.23	0.10	0.13	0.06	0.19	0.03	0.00	0.00	0.10
Omy_114976-223		0.04	0.03	0.04	0.09	0.00	0.07	0.17	0.42	0.35	0.37	0.48	0.03	0.44	0.33	0.44	0.28	0.06
Omy_115987-812		0.07	0.00	0.21	0.00	0.06	0.38	0.27	0.35	0.38	0.43	0.26	0.03	0.50	0.07	0.09	0.22	0.06
Omy_116104-229		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.05	0.20	0.15	0.00	0.17	0.00	0.00	0.00	0.06
Omy_116362-467		0.00	0.50	0.02	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.38	0.00	0.03	0.69	0.13	0.21
Omy_116733-349		0.39	0.00	0.21	0.47	0.03	0.44	0.88	0.86	0.59	0.63	0.81	0.09	0.56	0.10	0.38	0.13	0.59
Omy_116938-264		0.02	0.90	0.07	0.09	0.00	0.26	0.00	0.00	0.00	0.00	0.00	0.22	0.00	0.10	0.69	0.20	0.04
Omy_117242-419		0.34	0.83	0.46	0.19	0.84	0.57	0.41	0.57	0.40	0.57	0.44	0.41	0.33	1.00	0.28	0.74	0.50
Omy_117259-96		0.26	0.87	0.09	0.22	0.00	0.07	0.23	0.37	0.31	0.37	0.33	0.00	0.31	0.00	0.06	0.09	0.17
Omy_117286-374		0.04	0.00	0.02	0.00	0.38	0.09	0.02	0.21	0.26	0.20	0.25	0.06	0.17	0.17	0.03	0.07	0.38
Omy_117370-400		0.37	0.94	0.35	0.75	0.09	0.44	0.33	0.55	0.47	0.30	0.69	0.75	0.44	0.67	0.47	0.67	0.56
Omy_117432-190		0.00	0.53	0.11	0.81	0.00	0.04	0.13	0.04	0.08	0.00	0.00	0.50	0.00	0.23	0.00	0.00	0.00
Omy_117540-259		0.02	0.00	0.21	0.38	0.03	0.28	0.44	0.20	0.59	0.67	0.40	0.09	0.67	0.30	0.16	0.18	0.90
Omy_117549-316		0.00	0.87	0.02	0.06	0.66	0.66	0.19	0.38	0.26	0.30	0.37	0.69	0.33	0.13	0.83	0.70	0.54
Omy_117743-127		0.00	-	0.00	-	-	0.13	0.00	0.00	0.00	0.00	0.04	-	0.00	-	-	0.26	0.08
Omy_117815-81		0.33	0.00	0.46	0.13	0.38	0.36	0.19	0.09	0.15	0.13	0.11	0.06	0.06	0.10	0.38	0.07	0.00
Omy_118175-396		0.00	0.60	0.13	0.13	0.38	0.13	0.35	0.15	0.20	0.30	0.46	0.09	0.33	0.00	0.00	0.15	0.02
Omy_118205-116		0.04	0.50	0.37	0.13	0.09	0.30	0.61	0.49	0.51	0.47	0.48	0.66	0.31	0.10	0.34	0.28	0.94
Omy_118654-91		0.46	0.97	0.13	0.84	0.00	0.52	0.29	0.51	0.73	0.57	0.58	0.56	0.50	0.13	0.50	0.26	0.74
Omy_118938-341		0.00	0.17	0.65	0.00	0.16	0.41	0.10	0.21	0.09	0.03	0.23	0.00	0.06	0.07	0.00	0.02	0.23
Omy_119108-357		0.00	0.07	0.07	0.03	0.00	0.24	0.10	0.42	0.32	0.37	0.63	0.22	0.31	0.00	0.50	0.20	0.08
Omy_119892-365		0.13	0.00	0.33	0.03	0.31	0.20	0.33	0.18	0.12	0.17	0.15	0.13	0.28	0.00	0.41	0.41	0.06
Omy_120255-332		0.20	0.50	0.08	0.06	0.19	0.13	0.19	0.26	0.27	0.07	0.37	0.31	0.00	0.00	0.16	0.50	0.19
Omy_120950-569		0.38	0.63	0.16	0.38	0.09	0.36	0.48	0.16	0.12	0.07	0.06	0.81	0.06	0.13	0.44	0.35	0.29
Omy_121006-131		0.00	0.00	0.28	0.00	0.00	0.39	0.63	0.43	0.78	0.90	0.96	0.06	0.83	0.00	0.63	0.70	0.88

Table 1.3 Continued

		Willamette River, Wiley Creek	Klamath River, Buckboard Creek	Klamath River, Kelsey Creek	Goose Lake, Bauers Creek	McCloud River, Butcherknife Creek	Sacramento River, Battle Creek	Eel River, Middle Fork (summer)	Russian River	Scott Creek	San Lorenzo River	Carmel River	Eagle Lake Strain, A_RH	San Lorenzo Strain, BCH	Kamloops Strain, HCH	Mount Whitney Strain, FH	Lake Taupo, New Zealand	Río Santa Cruz, Argentina
	N	23	15	23	16	16	24	24	47	46	15	24	16	9	15	16	23	24
Assay Name		MAF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF	AF
Omy_121713-115		0.50	0.03	0.17	0.97	0.00	0.26	0.15	0.12	0.10	0.20	0.08	0.19	0.22	0.77	0.28	0.20	0.10
Omy_123044-128		0.00	0.10	0.17	0.34	0.03	0.28	0.29	0.46	0.54	0.50	0.46	0.16	0.38	0.00	0.31	0.76	0.87
Omy_123048-119		0.00	0.37	0.00	0.06	0.00	0.37	0.00	0.01	0.04	0.00	0.00	0.16	0.00	0.13	0.41	0.38	0.20
Omy_123921-144		0.41	0.00	0.32	0.19	0.00	0.15	0.27	0.27	0.19	0.30	0.17	0.03	0.22	0.33	0.03	0.46	0.02
Omy_124774-530		0.06	0.07	0.11	0.25	0.53	0.74	0.60	0.34	0.55	0.57	0.48	0.72	0.39	0.20	0.38	0.60	0.69
Omy_125998-61		0.00	0.13	0.04	0.03	0.06	0.26	0.56	0.66	0.70	0.73	0.85	0.22	0.72	0.23	0.50	0.48	0.94
Omy_126160-242		0.00	-	0.00	-	0.00	0.02	0.13	0.18	0.36	0.27	0.35	-	0.39	-	-	0.00	0.02
Omy_127236-583		0.00	0.00	0.30	0.00	0.00	0.23	0.60	0.68	0.76	0.87	0.72	0.03	0.83	0.00	0.16	0.37	0.78
Omy_127510-920		0.15	0.30	0.24	0.56	0.13	0.52	0.38	0.39	0.22	0.23	0.08	0.91	0.28	0.50	0.50	0.20	0.06
Omy_127645-308		0.00	0.07	0.00	0.81	0.00	0.30	0.00	0.06	0.01	0.00	0.02	0.72	0.00	0.03	0.00	0.20	0.35
Omy_127760-385		0.00	0.13	0.00	0.03	0.06	0.10	0.23	0.06	0.18	0.17	0.17	0.06	0.06	0.00	0.22	0.07	0.02
Omy_128302-430		0.04	0.00	0.00	0.00	0.00	0.00	0.10	0.02	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Omy_128693-755		0.32	0.60	0.70	0.00	0.41	0.30	0.42	0.30	0.27	0.23	0.30	0.44	0.17	0.00	0.00	0.24	0.28
Omy_128851-273		0.00	0.10	0.00	0.00	0.00	0.11	0.25	0.30	0.22	0.47	0.06	0.00	0.50	0.00	0.34	0.00	0.00
Omy_128923-433		0.09	0.00	0.05	0.00	0.00	0.02	0.07	-	0.13	0.03	-	0.17	0.00	0.50	0.19	0.22	0.37
Omy_128996-481		0.48	0.53	0.50	0.31	0.75	0.50	0.13	0.11	0.19	0.14	0.17	0.28	0.17	0.00	0.13	0.33	0.37
Omy_129170-794		0.00	0.03	0.02	0.00	0.06	0.11	0.10	0.46	0.38	0.36	0.27	0.22	0.33	0.00	0.00	0.20	0.00
Omy_129870-756		0.35	0.03	0.39	0.19	0.25	0.76	0.58	0.53	0.76	0.70	0.56	0.44	0.83	0.17	0.75	0.76	0.65
Omy_130295-98		0.35	0.40	0.64	0.09	0.28	0.67	0.75	0.94	0.88	0.97	1.00	0.09	1.00	0.37	0.84	0.67	0.56
Omy_130524-160		0.39	0.89	0.59	0.31	0.91	0.33	0.52	0.37	0.47	0.57	0.31	0.72	0.56	0.37	0.75	0.00	0.00
Omy_130720-100		0.00	0.33	0.14	0.03	0.91	0.63	0.58	0.59	0.50	0.70	0.69	0.78	0.69	0.10	0.63	0.59	0.60
Omy_131460-646		0.02	0.03	0.02	0.00	0.00	0.09	0.54	0.48	0.38	0.60	0.33	0.19	0.50	0.13	0.03	0.28	0.22
Omy_131965-120		0.11	0.80	0.35	0.13	0.00	0.57	0.40	0.49	0.39	0.33	0.19	0.41	0.44	0.17	0.56	0.41	0.23
Polymorphic loci (%)		66.2	73.4	84.2	70.5	60.4	97.1	86.3	95.7	91.4	88.5	89.2	86.3	81.3	75.5	83.5	88.5	90.6

Out of the 139 SNPs, 43 were identified in introns and 95 in other regions (CDSs or UTRs) based on the alignment of the consensus with the RtGI EST sequences. One consensus sequence did not align with the corresponding RtGI EST sequence. A total of 71 consensus sequences matched a known gene from GeneBank. Only 22 SNPs were identified in UTRs and 14 in CDSs based on those "blasted" genes that presented a translation. Finally, five SNPs corresponded to non-synonymous and nine to synonymous mutations (Table 1.5).

Based on the screen for the most informative, I selected 82 SNPs from this study, six from Campbell *et al.* (2009), three from Aguilar & Garza (2008), one from Campbell (unpubl.), and three from DeKoning (unpubl.; Table 1.5).

Table 1.4: Expected (H_E) and observed (H_O) heterozygosity of 139 SNP assays in 11 natural-origin steelhead populations, four *O. mykiss* hatchery strains, and two introduced *O. mykiss* populations in the southern hemisphere.

		Willamette River, Wiley Creek	Klamath River, Buckboard Creek	Klamath River, Kelsey Creek	Goose Lake, Bauers Creek	McCloud River, Butcherknife Creek	Sacramento River, Battle Creek	Eel River, Middle Fork (summer)	Russian River	Scott Creek	San Lorenzo River	Carmel River	Eagle Lake Strain, A_RH	San Lorenzo Strain, BCH	Kamloops Strain, HCH	Mount Whitney Strain, FH	Lake Taupo, New Zealand	Río Santa Cruz, Argentina	
Assay Name	N	23	15	23	16	16	24	24	47	46	15	24	16	9	15	16	23	24	F_{ST}
OMGH1PROM1-SNP1	H_E	0.09	0.27	0.32	0.42	0.12	0.51	0.44	0.43	0.38	0.19	0.36	0.18	0.21	0.13	0.47	0.41	0.43	0.076
	H_O	0.09	0.00	0.30	0.31	0.13	0.52	0.46	0.32	0.38	0.20	0.21	0.19	0.22	0.13	0.31	0.36	0.43	
Omy_95318-147	H_E	0.50	0.00	0.16	0.51	0.00	0.25	0.12	0.48	0.51	0.51	0.51	0.24	0.52	0.33	0.51	0.46	0.37	0.078
	H_O	0.65	0.00	0.17	0.75	0.00	0.29	0.13	0.57	0.49	0.73	0.57	0.27	0.44	0.40	0.56	0.52	0.22	
Omy_95442-108	H_E	0.00	0.43	0.09	0.18	0.00	0.26	0.04	0.41	0.29	0.30	0.48	0.06	0.23	0.00	0.00	0.00	0.04	0.049
	H_O	0.00	0.47	0.09	0.19	0.00	0.30	0.04	0.28	0.26	0.36	0.39	0.06	0.25	0.00	0.00	0.00	0.04	
Omy_95489-423	H_E	0.00	0.40	0.05	0.12	0.31	0.49	0.50	0.47	0.45	0.48	0.50	0.18	0.50	0.24	0.23	0.51	0.47	0.241
	H_O	0.00	0.53	0.05	0.13	0.25	0.43	0.46	0.36	0.37	0.60	0.54	0.19	0.75	0.27	0.25	0.64	0.54	
Omy_96158-277	H_E	0.00	0.37	0.30	0.12	0.42	0.46	0.04	0.41	0.50	0.40	0.40	0.47	0.42	0.24	0.39	0.37	0.36	0.058
	H_O	0.00	0.33	0.36	0.13	0.44	0.61	0.04	0.47	0.55	0.40	0.38	0.56	0.56	0.27	0.50	0.30	0.38	
Omy_96222-125	H_E	0.17	0.19	0.49	0.27	0.35	0.20	0.28	0.27	0.38	0.29	0.48	0.00	0.29	0.13	0.06	0.00	0.38	0.164
	H_O	0.09	0.20	0.62	0.31	0.44	0.13	0.33	0.19	0.29	0.07	0.58	0.00	0.33	0.13	0.06	0.00	0.25	
Omy_96529-231	H_E	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.20	0.33	0.12	0.00	0.37	0.00	0.00	0.00	0.00	0.091
	H_O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.18	0.40	0.13	0.00	0.22	0.00	0.00	0.00	0.00	
Omy_96899-148	H_E	0.26	0.00	0.16	0.39	0.00	0.51	0.45	0.49	0.48	0.51	0.50	0.48	0.53	0.51	0.44	0.48	0.04	0.202
	H_O	0.30	0.00	0.17	0.38	0.00	0.64	0.42	0.54	0.56	0.47	0.46	0.50	0.56	0.47	0.63	0.48	0.04	
Omy_97077-73	H_E	0.23	0.00	0.09	0.06	0.00	0.41	0.51	0.30	0.38	0.37	0.49	0.00	0.29	0.00	0.44	0.20	0.00	0.138
	H_O	0.26	0.00	0.09	0.06	0.00	0.39	0.58	0.36	0.45	0.33	0.50	0.00	0.33	0.00	0.50	0.22	0.00	
Omy_97660-230	H_E	0.47	0.00	0.00	0.35	0.00	0.29	0.40	0.34	0.11	0.07	0.20	0.18	0.11	0.43	0.31	0.43	0.08	0.091
	H_O	0.45	0.00	0.00	0.44	0.00	0.26	0.29	0.30	0.12	0.07	0.22	0.19	0.11	0.47	0.25	0.32	0.08	
Omy_97865-196	H_E	0.00	0.00	0.00	0.12	0.00	0.16	0.00	0.06	0.11	0.00	0.04	0.48	0.00	0.00	0.40	0.00	0.00	0.451
	H_O	0.00	0.00	0.00	0.13	0.00	0.17	0.00	0.06	0.11	0.00	0.04	0.25	0.00	0.00	0.13	0.00	0.00	
Omy_97954-618	H_E	0.00	0.13	0.05	0.35	0.52	0.41	0.51	0.50	0.47	0.46	0.47	0.27	0.42	0.48	0.42	0.37	0.38	0.158
	H_O	0.00	0.13	0.05	0.19	0.50	0.30	0.50	0.51	0.35	0.27	0.38	0.31	0.33	0.33	0.56	0.39	0.42	
Omy_98188-405	H_E	0.35	0.07	0.09	0.12	0.06	0.32	0.12	0.35	0.48	0.35	0.50	0.31	0.47	0.43	0.12	0.30	0.00	0.267
	H_O	0.35	0.07	0.09	0.13	0.06	0.39	0.13	0.36	0.58	0.43	0.54	0.13	0.22	0.20	0.13	0.36	0.00	
Omy_98409-549	H_E	0.16	0.19	0.39	0.00	0.06	0.44	0.45	0.48	0.51	0.51	0.44	0.39	0.52	0.00	0.00	0.11	0.25	0.212
	H_O	0.17	0.20	0.35	0.00	0.06	0.43	0.58	0.54	0.50	0.60	0.46	0.38	0.44	0.00	0.00	0.00	0.29	
Omy_98683-165	H_E	0.09	0.50	0.00	0.18	0.00	0.49	0.51	0.50	0.45	0.50	0.38	0.39	0.52	0.24	0.50	0.48	0.12	0.237
	H_O	0.09	0.53	0.00	0.19	0.00	0.43	0.52	0.49	0.50	0.13	0.33	0.38	0.67	0.27	0.44	0.39	0.13	
Omy_99300-202	H_E	0.29	0.00	0.04	0.00	0.00	0.35	0.00	0.30	0.29	0.07	0.25	0.27	0.00	0.52	0.44	0.00	0.04	0.339
	H_O	0.26	0.00	0.04	0.00	0.00	0.43	0.00	0.23	0.34	0.07	0.29	0.31	0.00	0.47	0.38	0.00	0.04	
Omy_100771-63	H_E	0.21	0.24	0.00	0.00	0.06	0.50	0.28	0.50	0.38	0.51	0.40	0.48	0.50	0.13	0.00	0.23	0.31	0.217
	H_O	0.23	0.27	0.00	0.00	0.06	0.26	0.33	0.62	0.27	0.47	0.38	0.38	0.75	0.13	0.00	0.26	0.21	
Omy_100974-386	H_E	0.05	0.50	0.32	0.48	0.27	0.49	0.31	0.39	0.28	0.13	0.34	0.35	0.11	0.19	0.51	0.43	0.28	0.143
	H_O	0.05	0.53	0.29	0.25	0.31	0.43	0.29	0.28	0.24	0.13	0.33	0.44	0.11	0.20	0.44	0.52	0.33	
Omy_101119-554	H_E	0.00	0.24	0.13	0.06	0.44	0.45	0.00	0.00	0.00	0.07	0.00	0.42	0.11	0.13	0.12	0.50	0.45	0.159
	H_O	0.00	0.27	0.14	0.06	0.25	0.30	0.00	0.00	0.00	0.07	0.00	0.44	0.11	0.13	0.13	0.52	0.33	
Omy_101341-188	H_E	0.00	0.00	0.00	0.00	0.00	0.09	0.40	0.14	0.37	0.19	0.16	0.00	0.11	0.00	0.00	0.26	0.22	0.120
	H_O	0.00	0.00	0.00	0.00	0.00	0.09	0.38	0.15	0.35	0.20	0.17	0.00	0.11	0.00	0.00	0.22	0.25	
Omy_101554-306	H_E	0.04	0.24	0.09	0.23	0.00	0.35	0.47	0.50	0.50	0.46	0.49	0.44	0.50	0.40	0.51	0.35	0.08	0.287
	H_O	0.04	0.27	0.10	0.13	0.00	0.35	0.46	0.45	0.47	0.40	0.54	0.25	0.56	0.40	0.75	0.26	0.08	

Table 1.4 Continued

Assay Name	N	Willamette River, Wiley Creek	Klamath River, Buckboard Creek	Klamath River, Kelsey Creek	Goose Lake, Bauers Creek	McCloud River, Butcherknife Creek	Sacramento River, Battle Creek	Eel River, Middle Fork (summer)	Russian River	Scott Creek	San Lorenzo River	Carmel River	Eagle Lake Strain, A_RH	San Lorenzo Strain, BCH	Kamloops Strain, HCH	Mount Whitney Strain, FH	Lake Taupo, New Zealand	Río Santa Cruz, Argentina	F_{ST}
		23	15	23	16	16	24	24	47	46	15	24	16	9	15	16	23	24	
Omy_101704-329	H_E	0.00	0.00	0.00	0.00	0.00	0.43	0.00	0.16	0.00	0.00	0.00	0.12	0.00	0.24	0.51	0.51	0.08	0.256
	H_O	0.00	0.00	0.00	0.00	0.00	0.43	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.13	0.44	0.70	0.08	
Omy_101770-410	H_E	0.23	0.13	0.30	0.00	0.27	0.35	0.34	0.43	0.34	0.37	0.45	0.48	0.46	0.37	0.23	0.12	0.19	0.079
	H_O	0.26	0.13	0.36	0.00	0.19	0.35	0.33	0.40	0.33	0.20	0.33	0.50	0.63	0.20	0.25	0.13	0.21	
Omy_101832-195	H_E	0.32	0.33	0.23	0.42	0.06	0.41	0.42	0.51	0.50	0.51	0.50	0.18	0.40	0.51	0.06	0.35	0.48	0.256
	H_O	0.22	0.27	0.26	0.44	0.06	0.39	0.42	0.63	0.51	0.40	0.42	0.19	0.25	0.60	0.06	0.26	0.50	
Omy_101993-189	H_E	0.05	0.00	0.05	0.00	0.42	0.50	0.08	0.46	0.50	0.51	0.48	0.50	0.46	0.43	0.23	0.09	0.51	0.203
	H_O	0.05	0.00	0.05	0.00	0.44	0.42	0.08	0.45	0.48	0.67	0.50	0.44	0.63	0.47	0.13	0.09	0.58	
Omy_102213-204	H_E	0.04	0.33	0.06	0.23	0.23	0.43	0.22	0.21	0.22	0.20	0.12	0.37	0.00	0.07	0.44	0.44	0.44	0.230
	H_O	0.04	0.40	0.06	0.25	0.25	0.26	0.25	0.19	0.20	0.21	0.13	0.33	0.00	0.07	0.50	0.36	0.54	
Omy_102420-634	H_E	0.26	0.37	0.43	0.00	0.48	0.37	0.50	0.49	0.45	0.37	0.44	0.12	0.47	0.51	0.42	0.49	0.45	0.220
	H_O	0.22	0.33	0.26	0.00	0.50	0.39	0.54	0.47	0.45	0.33	0.38	0.13	0.22	0.53	0.44	0.43	0.50	
Omy_102457-423	H_E	0.17	0.33	0.50	0.27	0.00	0.41	0.47	0.51	0.34	0.52	0.50	0.51	0.53	0.37	0.51	0.50	0.36	0.418
	H_O	0.09	0.27	0.35	0.19	0.00	0.30	0.27	0.36	0.11	0.17	0.11	0.38	0.13	0.33	0.23	0.61	0.21	
Omy_102505-102	H_E	0.38	0.07	0.23	0.00	0.06	0.04	0.36	0.24	0.36	0.24	0.25	0.00	0.37	0.13	0.27	0.21	0.04	0.446
	H_O	0.41	0.07	0.17	0.00	0.06	0.04	0.29	0.28	0.37	0.13	0.29	0.00	0.44	0.13	0.31	0.23	0.04	
Omy_102510-682	H_E	0.09	0.29	0.23	0.00	0.00	0.35	0.22	0.43	0.13	0.24	0.38	0.12	0.21	0.00	0.44	0.41	0.00	0.051
	H_O	0.09	0.20	0.26	0.00	0.00	0.43	0.25	0.45	0.13	0.13	0.42	0.13	0.22	0.00	0.38	0.30	0.00	
Omy_102867-443	H_E	0.00	0.13	0.00	0.12	0.00	0.35	0.42	0.50	0.40	0.43	0.50	0.00	0.42	0.07	0.35	0.41	0.22	0.409
	H_O	0.00	0.13	0.00	0.13	0.00	0.35	0.42	0.62	0.45	0.47	0.33	0.00	0.33	0.07	0.44	0.13	0.25	
Omy_103350-395	H_E	0.16	0.46	0.30	0.12	0.00	0.46	0.48	0.49	0.41	0.51	0.51	0.37	0.52	0.29	0.50	0.51	0.51	0.125
	H_O	0.17	0.27	0.27	0.13	0.00	0.43	0.50	0.45	0.35	0.53	0.61	0.33	0.67	0.20	0.81	0.45	0.39	
Omy_103577-379	H_E	0.31	0.33	0.09	0.00	0.00	0.12	0.16	0.50	0.42	0.50	0.45	0.06	0.50	0.00	0.44	0.35	0.19	0.466
	H_O	0.26	0.40	0.09	0.00	0.00	0.13	0.17	0.40	0.26	0.53	0.30	0.06	0.56	0.00	0.50	0.26	0.21	
Omy_103705-558	H_E	0.37	0.07	0.26	0.42	0.35	0.23	0.38	0.48	0.45	0.40	0.50	0.18	0.52	0.33	0.00	0.50	0.19	0.120
	H_O	0.30	0.07	0.30	0.31	0.31	0.26	0.42	0.38	0.40	0.40	0.38	0.06	0.67	0.40	0.00	0.70	0.21	
Omy_103713-53	H_E	0.00	0.48	0.04	0.00	0.18	0.51	0.36	0.50	0.50	0.51	0.36	0.27	0.50	0.30	0.39	0.46	0.38	0.178
	H_O	0.00	0.60	0.04	0.00	0.19	0.42	0.29	0.39	0.50	0.40	0.46	0.31	0.78	0.21	0.38	0.67	0.33	
Omy_104519-624	H_E	0.41	0.19	0.39	0.12	0.06	0.45	0.31	0.50	0.40	0.40	0.36	0.48	0.50	0.51	0.48	0.50	0.51	0.077
	H_O	0.39	0.20	0.43	0.13	0.06	0.57	0.29	0.50	0.33	0.40	0.29	0.38	0.56	0.40	0.50	0.39	0.54	
Omy_104569-114	H_E	0.43	0.33	0.51	0.00	0.00	0.41	0.42	0.41	0.26	0.29	0.41	0.00	0.42	0.07	0.24	0.23	0.40	0.107
	H_O	0.35	0.27	0.50	0.00	0.00	0.39	0.33	0.23	0.26	0.20	0.39	0.00	0.56	0.07	0.27	0.26	0.13	
Omy_105075-162	H_E	0.39	0.00	0.49	0.23	0.51	0.26	0.19	0.18	0.44	0.46	0.50	0.00	0.50	0.07	0.00	0.00	0.36	0.172
	H_O	0.35	0.00	0.55	0.13	0.56	0.22	0.21	0.20	0.50	0.40	0.71	0.00	0.33	0.07	0.00	0.00	0.46	
Omy_105105-448	H_E	0.46	0.19	0.43	0.31	0.27	0.37	0.49	0.28	0.15	0.00	0.34	0.12	0.00	0.51	0.51	0.00	0.47	0.108
	H_O	0.17	0.20	0.41	0.38	0.31	0.39	0.46	0.24	0.07	0.00	0.42	0.13	0.00	0.47	0.69	0.00	0.46	
Omy_105115-367	H_E	0.04	0.24	0.33	0.00	0.06	0.22	0.36	0.38	0.42	0.19	0.16	0.27	0.11	0.24	0.13	0.26	0.48	0.208
	H_O	0.04	0.27	0.41	0.00	0.06	0.25	0.38	0.30	0.44	0.20	0.17	0.31	0.11	0.13	0.13	0.22	0.50	
Omy_105235-713	H_E	0.00	0.00	0.00	0.00	0.00	0.12	0.34	0.25	0.37	0.19	0.22	0.00	0.11	0.00	0.00	0.38	0.28	0.333
	H_O	0.00	0.00	0.00	0.00	0.00	0.13	0.33	0.20	0.38	0.20	0.17	0.00	0.11	0.00	0.00	0.32	0.33	
Omy_105385-406	H_E	0.49	0.51	0.49	0.50	0.47	0.48	0.43	0.44	0.34	0.40	0.49	0.48	0.40	0.43	0.39	0.06	0.42	0.221
	H_O	0.61	0.47	0.32	0.56	0.44	0.55	0.26	0.43	0.29	0.13	0.38	0.38	0.50	0.47	0.50	0.06	0.42	
Omy_105386-347	H_E	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.35	0.45	0.29	0.48	0.47	0.21	0.00	0.00	0.16	0.00	0.351
	H_O	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.36	0.44	0.33	0.50	0.19	0.22	0.00	0.00	0.17	0.00	
Omy_105401-363	H_E	0.00	0.00	0.05	0.00	0.35	0.29	0.47	0.48	0.50	0.52	0.50	0.00	0.53	0.00	0.00	0.33	0.25	0.248

Table 1.4 Continued

Assay Name		Willamette River, Wiley Creek	Klamath River, Buckboard Creek	Klamath River, Kelsey Creek	Goose Lake, Bauers Creek	McCloud River, Butcherknife Creek	Sacramento River, Battle Creek	Eel River, Middle Fork (summer)	Russian River	Scott Creek	San Lorenzo River	Carmel River	Eagle Lake Strain, A_RH	San Lorenzo Strain, BCH	Kamloops Strain, HCH	Mount Whitney Strain, FH	Lake Taupo, New Zealand	Río Santa Cruz, Argentina	
	N	23	15	23	16	16	24	24	47	46	15	24	16	9	15	16	23	24	F_{ST}
	H_O	0.00	0.00	0.05	0.00	0.19	0.26	0.46	0.53	0.49	0.60	0.67	0.00	1.00	0.00	0.00	0.32	0.29	
Omy_105407-74	H_E	0.00	0.33	0.50	0.50	0.06	0.49	0.50	0.39	0.47	0.43	0.40	0.48	0.52	0.00	0.52	0.45	0.51	0.463
	H_O	0.00	0.40	0.48	0.44	0.06	0.78	0.58	0.36	0.33	0.33	0.54	0.50	0.44	0.00	0.50	0.65	0.83	
Omy_105714-265	H_E	0.09	0.24	0.26	0.51	0.50	0.49	0.38	0.27	0.42	0.48	0.36	0.31	0.53	0.29	0.39	0.39	0.48	0.125
	H_O	0.09	0.27	0.22	0.38	0.69	0.35	0.42	0.06	0.38	0.60	0.38	0.38	0.11	0.33	0.38	0.43	0.39	
Omy_105897-101	H_E	0.00	0.29	0.04	0.47	0.00	0.34	0.00	0.02	0.02	0.00	0.04	0.18	0.00	0.19	0.18	0.29	0.16	0.177
	H_O	0.00	0.20	0.04	0.31	0.00	0.33	0.00	0.02	0.02	0.00	0.04	0.19	0.00	0.20	0.19	0.17	0.17	
Omy_106172-332	H_E	0.00	0.46	0.46	0.00	0.48	0.23	0.16	0.29	0.00	0.07	0.00	0.00	0.00	0.07	0.23	0.16	0.22	0.153
	H_O	0.00	0.40	0.26	0.00	0.38	0.26	0.17	0.16	0.00	0.07	0.00	0.00	0.00	0.07	0.25	0.09	0.25	
Omy_106313-445	H_E	0.00	0.00	0.46	0.00	0.48	0.48	0.28	0.51	0.48	0.52	0.34	0.06	0.50	0.24	0.50	0.38	0.51	0.192
	H_O	0.00	0.00	0.52	0.00	0.75	0.39	0.33	0.49	0.51	0.60	0.33	0.06	0.56	0.13	0.44	0.41	0.52	
Omy_106560-58	H_E	0.00	0.13	0.04	0.23	0.51	0.46	0.31	0.32	0.35	0.07	0.32	0.50	0.00	0.07	0.35	0.51	0.29	0.145
	H_O	0.00	0.13	0.04	0.13	0.56	0.43	0.29	0.40	0.36	0.07	0.39	0.56	0.00	0.07	0.44	0.33	0.35	
Omy_106747-707	H_E	0.51	0.40	0.51	0.27	0.51	0.49	0.51	0.48	0.44	0.52	0.48	0.51	0.53	0.37	0.51	0.51	0.49	0.179
	H_O	0.61	0.40	0.52	0.06	0.63	0.32	0.75	0.32	0.45	0.50	0.50	0.38	0.63	0.47	0.38	0.36	0.43	
Omy_107031-704	H_E	0.49	0.40	0.20	0.42	0.44	0.48	0.16	0.06	0.09	0.00	0.00	0.23	0.00	0.24	0.48	0.29	0.51	0.277
	H_O	0.73	0.53	0.13	0.44	0.50	0.42	0.17	0.02	0.00	0.00	0.00	0.25	0.00	0.27	0.50	0.35	0.46	
Omy_107074-217	H_E	0.00	0.00	0.04	0.00	0.00	0.26	0.51	0.42	0.20	0.19	0.34	0.12	0.21	0.00	0.47	0.37	0.08	0.153
	H_O	0.00	0.00	0.04	0.00	0.00	0.30	0.58	0.30	0.13	0.20	0.17	0.13	0.22	0.00	0.56	0.39	0.08	
Omy_107285-69	H_E	0.46	0.40	0.16	0.31	0.00	0.50	0.34	0.41	0.40	0.46	0.45	0.23	0.47	0.37	0.51	0.51	0.48	0.330
	H_O	0.43	0.40	0.17	0.25	0.00	0.52	0.42	0.46	0.40	0.67	0.42	0.25	0.44	0.33	0.69	0.48	0.57	
Omy_107336-170	H_E	0.09	0.13	0.20	0.52	0.00	0.37	0.00	0.02	0.00	0.00	0.00	0.42	0.00	0.40	0.00	0.49	0.16	0.348
	H_O	0.09	0.13	0.22	0.75	0.00	0.48	0.00	0.02	0.00	0.00	0.00	0.44	0.00	0.27	0.00	0.52	0.17	
Omy_107607-137	H_E	0.51	0.13	0.05	0.00	0.35	0.23	0.48	0.24	0.31	0.29	0.16	0.06	0.37	0.29	0.27	0.12	0.00	0.176
	H_O	0.52	0.13	0.05	0.00	0.44	0.26	0.50	0.23	0.33	0.20	0.17	0.06	0.44	0.20	0.31	0.13	0.00	
Omy_107786-314	H_E	0.16	0.07	0.20	0.23	0.00	0.41	0.19	0.42	0.21	0.29	0.34	0.43	0.21	0.19	0.50	0.29	0.49	0.079
	H_O	0.17	0.07	0.22	0.25	0.00	0.39	0.21	0.23	0.15	0.33	0.42	0.33	0.22	0.07	0.56	0.26	0.38	
Omy_107786-584	H_E	0.05	0.00	0.21	0.23	0.00	0.21	0.20	0.34	0.16	0.20	0.21	0.44	0.11	0.52	0.39	0.16	0.48	0.191
	H_O	0.05	0.00	0.24	0.25	0.00	0.24	0.22	0.19	0.13	0.21	0.24	0.38	0.11	0.47	0.38	0.09	0.30	
Omy_107806-34	H_E	0.49	0.43	0.50	0.06	0.31	0.49	0.04	0.09	0.15	0.33	0.04	0.23	0.37	0.52	0.27	0.37	0.38	0.344
	H_O	0.30	0.33	0.59	0.06	0.38	0.43	0.04	0.05	0.02	0.40	0.04	0.25	0.22	0.47	0.19	0.30	0.50	
Omy_108007-193	H_E	0.37	0.13	0.21	0.44	0.51	0.46	0.51	0.19	0.02	0.13	0.08	0.39	0.11	0.37	0.52	0.50	0.44	0.202
	H_O	0.48	0.00	0.23	0.38	0.88	0.52	0.54	0.17	0.02	0.13	0.08	0.38	0.11	0.47	0.63	0.57	0.46	
Omy_108735-311	H_E	0.50	0.19	0.49	0.06	0.44	0.51	0.48	0.41	0.32	0.46	0.29	0.39	0.52	0.29	0.50	0.45	0.50	0.292
	H_O	0.48	0.20	0.47	0.06	0.38	0.70	0.42	0.39	0.26	0.40	0.35	0.25	0.44	0.33	0.44	0.39	0.50	
Omy_108820-85	H_E	0.04	0.19	0.23	0.06	0.47	0.26	0.00	0.02	0.00	0.00	0.00	0.46	0.00	0.19	0.27	0.35	0.16	0.153
	H_O	0.04	0.20	0.26	0.06	0.31	0.22	0.00	0.02	0.00	0.00	0.00	0.53	0.00	0.20	0.31	0.26	0.17	
Omy_109243-222	H_E	0.16	0.37	0.49	0.27	0.06	0.51	0.22	0.23	0.22	0.37	0.16	0.31	0.29	0.13	0.51	0.50	0.51	0.273
	H_O	0.09	0.20	0.26	0.19	0.06	0.43	0.25	0.21	0.24	0.33	0.17	0.25	0.33	0.13	0.44	0.57	0.46	
Omy_109390-341	H_E	0.00	0.19	0.00	0.00	0.00	0.39	0.41	0.23	0.32	0.21	0.34	0.31	0.11	0.37	0.51	0.26	0.23	0.357
	H_O	0.00	0.20	0.00	0.00	0.00	0.52	0.30	0.26	0.26	0.23	0.42	0.38	0.11	0.33	0.56	0.30	0.17	
Omy_109525-403	H_E	0.51	0.13	0.47	0.12	0.12	0.39	0.40	0.50	0.48	0.35	0.38	0.18	0.42	0.52	0.51	0.32	0.22	0.204
	H_O	0.50	0.13	0.40	0.13	0.13	0.43	0.38	0.50	0.49	0.43	0.33	0.19	0.33	0.60	0.50	0.30	0.25	
Omy_109651-445	H_E	0.00	0.00	0.00	0.06	0.00	0.04	0.04	0.30	0.37	0.29	0.42	0.00	0.37	0.00	0.51	0.16	0.22	0.090
	H_O	0.00	0.00	0.00	0.06	0.00	0.04	0.04	0.27	0.48	0.33	0.50	0.00	0.22	0.00	0.38	0.17	0.17	

Table 1.4 Continued

Assay Name	N	Willamette River, Wiley Creek	Klamath River, Buckboard Creek	Klamath River, Kelsey Creek	Goose Lake, Bauers Creek	McCloud River, Butcherknife Creek	Sacramento River, Battle Creek	Eel River, Middle Fork (summer)	Russian River	Scott Creek	San Lorenzo River	Carmel River	Eagle Lake Strain, A_RH	San Lorenzo Strain, BCH	Kamloops Strain, HCH	Mount Whitney Strain, FH	Lake Taupo, New Zealand	Río Santa Cruz, Argentina	
		23	15	23	16	16	24	24	47	46	15	24	16	9	15	16	23	24	F_{ST}
Omy_109693-461	H_E	0.04	0.51	0.16	0.27	0.00	0.16	0.51	0.51	0.49	0.39	0.45	0.48	0.36	0.00	0.51	0.20	0.51	0.182
	H_O	0.04	0.40	0.17	0.31	0.00	0.17	0.38	0.55	0.48	0.50	0.57	0.63	0.43	0.00	0.38	0.13	0.65	
Omy_109874-148	H_E	0.12	0.00	0.00	0.00	0.00	0.04	0.16	0.24	0.37	0.50	0.38	0.00	0.52	0.00	0.18	0.20	0.00	0.162
	H_O	0.13	0.00	0.00	0.00	0.00	0.04	0.17	0.23	0.36	0.40	0.33	0.00	0.44	0.00	0.19	0.22	0.00	
Omy_109894-185	H_E	0.37	0.49	0.41	0.18	0.18	0.51	0.22	0.32	0.24	0.07	0.34	0.47	0.00	0.51	0.18	0.18	0.42	0.079
	H_O	0.38	0.00	0.11	0.19	0.19	0.39	0.17	0.39	0.27	0.07	0.08	0.69	0.00	0.60	0.19	0.20	0.33	
Omy_109944-74	H_E	0.00	0.00	0.12	0.12	0.00	0.49	0.28	0.19	0.07	0.07	0.04	0.27	0.00	0.00	0.27	0.39	0.12	0.073
	H_O	0.00	0.00	0.13	0.00	0.00	0.36	0.25	0.09	0.07	0.07	0.04	0.31	0.00	0.00	0.31	0.26	0.13	
Omy_110064-419	H_E	0.00	0.40	0.32	0.51	0.27	0.43	0.17	0.20	0.42	0.51	0.42	0.18	0.53	0.37	0.48	0.41	0.48	0.077
	H_O	0.00	0.40	0.30	0.63	0.19	0.52	0.18	0.13	0.33	0.67	0.42	0.19	0.56	0.33	0.25	0.39	0.42	
Omy_110078-294	H_E	0.20	0.29	0.35	0.50	0.00	0.51	0.19	0.38	0.44	0.29	0.25	0.12	0.29	0.33	0.43	0.44	0.40	0.096
	H_O	0.22	0.33	0.35	0.44	0.00	0.48	0.21	0.34	0.37	0.33	0.29	0.13	0.11	0.40	0.33	0.45	0.38	
Omy_110201-359	H_E	0.43	0.29	0.51	0.00	0.23	0.32	0.51	0.50	0.51	0.50	0.48	0.31	0.50	0.19	0.47	0.51	0.50	0.100
	H_O	0.52	0.20	0.43	0.00	0.25	0.39	0.50	0.53	0.41	0.53	0.42	0.38	0.50	0.20	0.44	0.48	0.58	
Omy_110362-585	H_E	0.20	0.00	0.09	0.31	0.12	0.49	0.50	0.50	0.44	0.51	0.51	0.23	0.52	0.33	0.00	0.26	0.48	0.169
	H_O	0.04	0.00	0.09	0.38	0.13	0.61	0.46	0.34	0.48	0.93	0.58	0.25	0.67	0.40	0.00	0.30	0.58	
Omy_110571-386	H_E	0.16	0.00	0.41	0.00	0.12	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.43	0.18	0.00	0.12	0.142
	H_O	0.17	0.00	0.39	0.00	0.13	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.60	0.19	0.00	0.13	
Omy_110689-148	H_E	0.51	0.33	0.51	0.51	0.06	0.49	0.51	0.48	0.51	0.51	0.51	0.44	0.42	0.29	0.51	0.51	0.45	0.163
	H_O	0.39	0.13	0.41	0.50	0.06	0.52	0.42	0.57	0.53	0.60	0.58	0.50	0.33	0.33	0.50	0.55	0.67	
Omy_111005-159	H_E	0.49	0.13	0.21	0.42	0.00	0.50	0.36	0.25	0.02	0.13	0.28	0.50	0.00	0.00	0.42	0.26	0.51	0.176
	H_O	0.61	0.13	0.24	0.31	0.00	0.52	0.46	0.24	0.02	0.13	0.33	0.69	0.00	0.00	0.31	0.22	0.46	
Omy_111084-526	H_E	0.23	0.00	0.09	0.00	0.00	0.26	0.04	0.08	0.20	0.19	0.31	0.48	0.21	0.19	0.00	0.00	0.34	0.299
	H_O	0.26	0.00	0.09	0.00	0.00	0.22	0.04	0.09	0.22	0.20	0.29	0.38	0.22	0.20	0.00	0.00	0.33	
Omy_111383-51	H_E	0.41	0.07	0.49	0.39	0.39	0.50	0.42	0.26	0.26	0.24	0.12	0.51	0.29	0.48	0.42	0.49	0.47	0.351
	H_O	0.30	0.07	0.55	0.25	0.38	0.61	0.50	0.26	0.26	0.27	0.13	0.31	0.33	0.73	0.31	0.35	0.54	
Omy_111666-301	H_E	0.04	0.19	0.41	0.00	0.23	0.51	0.50	0.50	0.50	0.52	0.51	0.52	0.50	0.37	0.51	0.45	0.50	0.567
	H_O	0.04	0.07	0.30	0.00	0.13	0.48	0.50	0.57	0.56	0.73	0.38	0.50	0.25	0.47	0.25	0.39	0.50	
Omy_111681-432	H_E	0.00	0.07	0.09	0.23	0.00	0.26	0.00	0.00	0.00	0.00	0.00	0.39	0.00	0.07	0.06	0.48	0.47	0.178
	H_O	0.00	0.07	0.09	0.25	0.00	0.22	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.07	0.06	0.48	0.54	
Omy_112208-328	H_E	0.44	0.46	0.51	0.35	0.27	0.43	0.40	0.48	0.45	0.50	0.50	0.31	0.52	0.33	0.39	0.51	0.50	0.203
	H_O	0.39	0.40	0.48	0.31	0.31	0.52	0.46	0.30	0.64	0.53	0.52	0.25	0.67	0.40	0.38	0.61	0.43	
Omy_112301-202	H_E	0.17	0.52	0.47	0.48	0.31	0.39	0.38	0.04	0.10	0.00	0.00	0.44	0.00	0.24	0.51	0.26	0.34	0.270
	H_O	0.18	0.47	0.45	0.50	0.38	0.26	0.42	0.04	0.02	0.00	0.00	0.50	0.00	0.27	0.69	0.30	0.42	
Omy_112820-82	H_E	0.09	0.00	0.42	0.18	0.35	0.50	0.25	0.47	0.30	0.29	0.00	0.29	0.37	0.13	0.39	0.51	0.51	0.116
	H_O	0.09	0.00	0.48	0.19	0.44	0.39	0.29	0.38	0.27	0.33	0.00	0.20	0.22	0.13	0.38	0.39	0.42	
Omy_112876-45	H_E	0.51	0.51	0.35	0.44	0.42	0.39	0.50	0.50	0.51	0.51	0.49	0.35	0.53	0.50	0.39	0.39	0.50	0.206
	H_O	0.86	0.47	0.43	0.38	0.44	0.52	0.88	0.96	0.96	0.87	0.79	0.31	1.00	0.80	0.50	0.52	0.83	
Omy_113109-205	H_E	0.00	0.00	0.00	0.00	0.00	0.45	0.08	0.50	0.42	0.50	0.40	0.39	0.52	0.19	0.51	0.35	0.34	0.213
	H_O	0.00	0.00	0.00	0.00	0.00	0.39	0.08	0.35	0.47	0.53	0.38	0.21	0.67	0.20	0.50	0.26	0.42	
Omy_113128-73	H_E	0.00	0.00	0.26	0.06	0.06	0.12	0.45	0.50	0.50	0.52	0.48	0.18	0.50	0.07	0.12	0.39	0.19	0.091
	H_O	0.00	0.00	0.22	0.06	0.06	0.13	0.42	0.33	0.47	0.47	0.57	0.19	0.50	0.07	0.13	0.35	0.13	
Omy_113242-163	H_E	0.00	0.33	0.12	0.31	0.00	0.16	0.00	0.02	0.09	0.00	0.00	0.50	0.00	0.07	0.00	0.20	0.04	0.193
	H_O	0.00	0.27	0.13	0.25	0.00	0.17	0.00	0.02	0.00	0.00	0.00	0.31	0.00	0.07	0.00	0.22	0.04	
Omy_113490-159	H_E	0.47	0.43	0.51	0.31	0.06	0.29	0.28	0.02	0.00	0.00	0.00	0.47	0.00	0.48	0.50	0.04	0.50	0.212

Table 1.4 Continued

Assay Name		Willamette River, Wiley Creek	Klamath River, Buckboard Creek	Klamath River, Kelsey Creek	Goose Lake, Bauers Creek	McCloud River, Butcherknife Creek	Sacramento River, Battle Creek	Eel River, Middle Fork (summer)	Russian River	Scott Creek	San Lorenzo River	Carmel River	Eagle Lake Strain, A_RH	San Lorenzo Strain, BCH	Kamloops Strain, HCH	Mount Whitney Strain, FH	Lake Taupo, New Zealand	Río Santa Cruz, Argentina	
	N	23	15	23	16	16	24	24	47	46	15	24	16	9	15	16	23	24	F_{ST}
	H_O	0.64	0.47	0.50	0.38	0.06	0.35	0.33	0.02	0.00	0.00	0.00	0.44	0.00	0.47	0.44	0.04	0.58	
Omy_114315-438	H_E	0.09	0.13	0.26	0.27	0.48	0.46	0.22	0.45	0.49	0.43	0.45	0.12	0.37	0.24	0.27	0.00	0.22	0.122
	H_O	0.09	0.13	0.30	0.31	0.50	0.52	0.25	0.32	0.47	0.60	0.39	0.00	0.44	0.27	0.31	0.00	0.17	
Omy_114448-87	H_E	0.00	0.00	0.44	0.00	0.00	0.50	0.44	0.49	0.30	0.19	0.08	0.12	0.29	0.00	0.51	0.50	0.50	0.210
	H_O	0.00	0.00	0.55	0.00	0.00	0.57	0.46	0.46	0.28	0.20	0.08	0.13	0.33	0.00	0.63	0.61	0.46	
Omy_114587-480	H_E	0.29	0.19	0.13	0.23	0.00	0.04	0.00	0.31	0.36	0.19	0.22	0.12	0.33	0.07	0.00	0.00	0.19	0.137
	H_O	0.26	0.20	0.14	0.25	0.00	0.04	0.00	0.26	0.33	0.20	0.25	0.13	0.13	0.07	0.00	0.00	0.21	
Omy_114976-223	H_E	0.09	0.07	0.09	0.18	0.00	0.12	0.28	0.49	0.46	0.48	0.51	0.06	0.53	0.46	0.51	0.41	0.12	0.213
	H_O	0.09	0.07	0.09	0.19	0.00	0.13	0.33	0.40	0.43	0.47	0.63	0.06	0.13	0.53	0.63	0.30	0.13	
Omy_115987-812	H_E	0.12	0.00	0.33	0.00	0.12	0.48	0.40	0.46	0.48	0.51	0.39	0.06	0.53	0.13	0.18	0.35	0.12	0.102
	H_O	0.13	0.00	0.41	0.00	0.13	0.67	0.54	0.45	0.37	0.73	0.35	0.06	0.56	0.13	0.19	0.35	0.13	
Omy_116104-229	H_E	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.09	0.33	0.25	0.00	0.29	0.00	0.00	0.00	0.12	0.210
	H_O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.09	0.40	0.29	0.00	0.33	0.00	0.00	0.00	0.13	
Omy_116362-467	H_E	0.00	0.52	0.04	0.00	0.00	0.37	0.00	0.00	0.00	0.00	0.00	0.48	0.00	0.07	0.44	0.23	0.33	0.192
	H_O	0.00	0.60	0.04	0.00	0.00	0.39	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.07	0.50	0.26	0.41	
Omy_116733-349	H_E	0.49	0.00	0.33	0.51	0.06	0.50	0.22	0.24	0.49	0.48	0.31	0.18	0.52	0.19	0.48	0.23	0.50	0.214
	H_O	0.43	0.00	0.23	0.56	0.06	0.52	0.25	0.19	0.52	0.47	0.29	0.19	0.44	0.20	0.38	0.17	0.48	
Omy_116938-264	H_E	0.04	0.19	0.12	0.18	0.00	0.39	0.00	0.00	0.00	0.00	0.00	0.35	0.00	0.19	0.44	0.32	0.08	-0.011
	H_O	0.04	0.20	0.13	0.19	0.00	0.43	0.00	0.00	0.00	0.00	0.00	0.44	0.00	0.20	0.38	0.30	0.00	
Omy_117242-419	H_E	0.46	0.29	0.51	0.31	0.27	0.50	0.50	0.49	0.48	0.51	0.50	0.50	0.47	0.00	0.42	0.39	0.51	0.113
	H_O	0.50	0.33	0.39	0.25	0.31	0.26	0.48	0.34	0.57	0.43	0.54	0.56	0.22	0.00	0.44	0.43	0.58	
Omy_117259-96	H_E	0.39	0.24	0.16	0.35	0.00	0.12	0.36	0.47	0.43	0.48	0.45	0.00	0.46	0.00	0.12	0.16	0.28	0.210
	H_O	0.43	0.27	0.17	0.44	0.00	0.13	0.38	0.40	0.48	0.47	0.42	0.00	0.63	0.00	0.13	0.17	0.33	
Omy_117286-374	H_E	0.09	0.00	0.04	0.00	0.48	0.16	0.04	0.33	0.38	0.33	0.38	0.12	0.29	0.29	0.06	0.12	0.48	0.249
	H_O	0.09	0.00	0.04	0.00	0.38	0.17	0.04	0.37	0.42	0.27	0.42	0.13	0.33	0.33	0.06	0.13	0.67	
Omy_117370-400	H_E	0.48	0.13	0.46	0.39	0.18	0.50	0.45	0.50	0.50	0.43	0.44	0.39	0.53	0.46	0.51	0.45	0.50	0.126
	H_O	0.39	0.13	0.43	0.38	0.19	0.43	0.33	0.43	0.30	0.60	0.29	0.38	0.63	0.53	0.44	0.39	0.63	
Omy_117432-190	H_E	0.00	0.51	0.20	0.31	0.00	0.09	0.22	0.08	0.14	0.00	0.00	0.52	0.00	0.37	0.00	0.00	0.00	0.123
	H_O	0.00	0.53	0.22	0.38	0.00	0.09	0.25	0.09	0.02	0.00	0.00	0.38	0.00	0.33	0.00	0.00	0.00	
Omy_117540-259	H_E	0.04	0.00	0.34	0.48	0.06	0.41	0.50	0.33	0.49	0.46	0.49	0.18	0.47	0.43	0.27	0.30	0.19	0.230
	H_O	0.04	0.00	0.33	0.50	0.06	0.39	0.43	0.28	0.56	0.53	0.63	0.19	0.44	0.07	0.31	0.36	0.21	
Omy_117549-316	H_E	0.00	0.24	0.04	0.12	0.47	0.46	0.31	0.48	0.39	0.43	0.48	0.44	0.47	0.24	0.29	0.43	0.51	0.258
	H_O	0.00	0.27	0.04	0.13	0.44	0.41	0.38	0.43	0.39	0.33	0.39	0.38	0.44	0.27	0.20	0.43	0.65	
Omy_117743-127	H_E	0.00	-	0.00	-	-	0.23	0.00	0.00	0.00	0.00	0.08	-	0.00	-	-	0.39	0.16	0.245
	H_O	0.00	-	0.00	-	-	0.26	0.00	0.00	0.00	0.00	0.08	-	0.00	-	-	0.35	0.17	
Omy_117815-81	H_E	0.45	0.00	0.51	0.23	0.48	0.47	0.31	0.16	0.26	0.24	0.20	0.12	0.11	0.19	0.48	0.13	0.00	0.120
	H_O	0.57	0.00	0.48	0.13	0.25	0.45	0.38	0.17	0.30	0.27	0.22	0.13	0.11	0.20	0.50	0.14	0.00	
Omy_118175-396	H_E	0.00	0.50	0.23	0.23	0.48	0.22	0.47	0.26	0.32	0.43	0.51	0.18	0.47	0.00	0.00	0.26	0.04	0.476
	H_O	0.00	0.67	0.17	0.25	0.38	0.25	0.54	0.26	0.30	0.60	0.42	0.06	0.67	0.00	0.00	0.30	0.04	
Omy_118205-116	H_E	0.09	0.52	0.48	0.23	0.18	0.43	0.49	0.51	0.51	0.51	0.51	0.47	0.46	0.19	0.47	0.41	0.12	0.067
	H_O	0.09	0.64	0.63	0.13	0.19	0.26	0.43	0.53	0.62	0.67	0.54	0.31	0.13	0.20	0.56	0.30	0.13	
Omy_118654-91	H_E	0.51	0.07	0.23	0.27	0.00	0.51	0.42	0.51	0.40	0.51	0.50	0.51	0.53	0.24	0.52	0.39	0.39	0.133
	H_O	0.57	0.07	0.26	0.06	0.00	0.43	0.50	0.51	0.40	0.57	0.58	0.38	0.56	0.27	0.38	0.35	0.35	
Omy_118938-341	H_E	0.00	0.29	0.47	0.00	0.27	0.50	0.19	0.33	0.16	0.07	0.36	0.00	0.11	0.13	0.00	0.04	0.36	0.096
	H_O	0.00	0.07	0.40	0.00	0.31	0.65	0.21	0.24	0.18	0.07	0.29	0.00	0.11	0.13	0.00	0.04	0.29	

Table 1.4 Continued

Assay Name		Willamette River, Wiley Creek	Klamath River, Buckboard Creek	Klamath River, Kelsey Creek	Goose Lake, Bauers Creek	McCloud River, Butcherknife Creek	Sacramento River, Battle Creek	Eel River, Middle Fork (summer)	Russian River	Scott Creek	San Lorenzo River	Carmel River	Eagle Lake Strain, A_RH	San Lorenzo Strain, BCH	Kamloops Strain, HCH	Mount Whitney Strain, FH	Lake Taupo, New Zealand	Río Santa Cruz, Argentina	
	N	23	15	23	16	16	24	24	47	46	15	24	16	9	15	16	23	24	F_{ST}
Omy_119108-357	H_E	0.00	0.14	0.12	0.06	0.00	0.37	0.19	0.49	0.44	0.48	0.48	0.35	0.46	0.00	0.52	0.32	0.16	0.117
	H_O	0.00	0.14	0.13	0.06	0.00	0.39	0.13	0.57	0.47	0.47	0.42	0.31	0.63	0.00	0.50	0.22	0.17	
Omy_119892-365	H_E	0.23	0.00	0.45	0.06	0.44	0.32	0.45	0.30	0.21	0.29	0.25	0.23	0.42	0.00	0.50	0.50	0.12	0.501
	H_O	0.26	0.00	0.55	0.06	0.63	0.30	0.48	0.32	0.19	0.20	0.29	0.25	0.33	0.00	0.56	0.48	0.13	
Omy_120255-332	H_E	0.32	0.53	0.14	0.12	0.31	0.23	0.31	0.38	0.40	0.13	0.48	0.44	0.00	0.00	0.27	0.51	0.31	0.102
	H_O	0.39	0.00	0.15	0.13	0.25	0.17	0.21	0.29	0.34	0.13	0.57	0.38	0.00	0.00	0.19	0.57	0.29	
Omy_120950-569	H_E	0.48	0.48	0.27	0.48	0.18	0.47	0.51	0.28	0.21	0.13	0.12	0.31	0.11	0.24	0.51	0.46	0.42	0.119
	H_O	0.25	0.47	0.32	0.50	0.19	0.36	0.43	0.15	0.19	0.13	0.13	0.38	0.11	0.13	0.63	0.52	0.25	
Omy_121006-131	H_E	0.00	0.00	0.41	0.00	0.00	0.49	0.48	0.49	0.34	0.19	0.09	0.12	0.29	0.00	0.48	0.43	0.22	0.193
	H_O	0.00	0.00	0.48	0.00	0.00	0.43	0.42	0.47	0.30	0.20	0.09	0.13	0.33	0.00	0.63	0.35	0.25	
Omy_121713-115	H_E	0.51	0.07	0.29	0.06	0.00	0.39	0.25	0.21	0.18	0.33	0.16	0.31	0.37	0.37	0.42	0.32	0.19	0.060
	H_O	0.11	0.07	0.35	0.06	0.00	0.52	0.29	0.23	0.16	0.27	0.17	0.38	0.44	0.20	0.19	0.30	0.21	
Omy_123044-128	H_E	0.00	0.19	0.28	0.47	0.06	0.41	0.42	0.50	0.50	0.52	0.51	0.27	0.50	0.00	0.44	0.37	0.23	0.117
	H_O	0.00	0.20	0.33	0.44	0.06	0.48	0.42	0.66	0.43	0.60	0.25	0.31	0.50	0.00	0.38	0.30	0.26	
Omy_123048-119	H_E	0.00	0.48	0.00	0.12	0.00	0.48	0.00	0.02	0.09	0.00	0.00	0.27	0.00	0.24	0.50	0.48	0.32	0.428
	H_O	0.00	0.47	0.00	0.13	0.00	0.48	0.00	0.02	0.00	0.00	0.00	0.06	0.00	0.27	0.56	0.48	0.30	
Omy_123921-144	H_E	0.50	0.00	0.44	0.31	0.00	0.26	0.40	0.40	0.32	0.43	0.28	0.06	0.37	0.46	0.06	0.51	0.04	0.368
	H_O	0.65	0.00	0.55	0.38	0.00	0.30	0.54	0.37	0.25	0.33	0.33	0.06	0.44	0.27	0.06	0.57	0.04	
Omy_124774-530	H_E	0.12	0.13	0.20	0.39	0.51	0.39	0.49	0.45	0.50	0.51	0.51	0.42	0.50	0.33	0.48	0.49	0.44	0.109
	H_O	0.13	0.13	0.22	0.25	0.53	0.35	0.71	0.36	0.62	0.60	0.63	0.56	0.33	0.27	0.50	0.71	0.46	
Omy_125998-61	H_E	0.00	0.24	0.09	0.06	0.12	0.39	0.50	0.45	0.43	0.40	0.25	0.35	0.42	0.37	0.52	0.51	0.12	0.028
	H_O	0.00	0.27	0.09	0.06	0.13	0.35	0.46	0.33	0.30	0.53	0.29	0.44	0.56	0.47	0.38	0.52	0.13	
Omy_126160-242	H_E	0.00	-	0.00	-	0.00	0.04	0.23	0.30	0.47	0.40	0.47	-	0.50	-	-	0.00	0.04	0.413
	H_O	0.00	-	0.00	-	0.00	0.04	0.26	0.23	0.59	0.40	0.46	-	0.33	-	-	0.00	0.04	
Omy_127236-583	H_E	0.00	0.00	0.43	0.00	0.00	0.36	0.49	0.44	0.37	0.24	0.41	0.06	0.29	0.00	0.27	0.48	0.35	0.443
	H_O	0.00	0.00	0.52	0.00	0.00	0.36	0.46	0.47	0.36	0.27	0.22	0.06	0.33	0.00	0.31	0.65	0.43	
Omy_127510-920	H_E	0.26	0.43	0.37	0.51	0.23	0.51	0.48	0.48	0.35	0.37	0.16	0.18	0.42	0.52	0.52	0.32	0.12	0.140
	H_O	0.13	0.47	0.39	0.38	0.25	0.52	0.50	0.32	0.31	0.47	0.17	0.19	0.33	0.33	0.50	0.39	0.04	
Omy_127645-308	H_E	0.00	0.13	0.00	0.31	0.00	0.43	0.00	0.12	0.02	0.00	0.04	0.42	0.00	0.07	0.00	0.32	0.46	0.209
	H_O	0.00	0.13	0.00	0.38	0.00	0.43	0.00	0.13	0.02	0.00	0.04	0.44	0.00	0.07	0.00	0.39	0.43	
Omy_127760-385	H_E	0.00	0.24	0.00	0.06	0.12	0.19	0.36	0.12	0.30	0.29	0.28	0.12	0.11	0.00	0.35	0.12	0.04	0.194
	H_O	0.00	0.27	0.00	0.06	0.13	0.21	0.38	0.13	0.31	0.33	0.33	0.13	0.11	0.00	0.31	0.13	0.04	
Omy_128302-430	H_E	0.09	0.00	0.00	0.00	0.00	0.00	0.19	0.04	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.307
	H_O	0.09	0.00	0.00	0.00	0.00	0.00	0.21	0.04	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Omy_128693-755	H_E	0.44	0.50	0.43	0.00	0.50	0.43	0.50	0.42	0.40	0.37	0.43	0.51	0.29	0.00	0.00	0.37	0.41	0.181
	H_O	0.32	0.40	0.40	0.00	0.69	0.60	0.33	0.33	0.44	0.20	0.60	0.63	0.33	0.00	0.00	0.48	0.35	
Omy_128851-273	H_E	0.00	0.19	0.00	0.00	0.00	0.20	0.38	0.42	0.34	0.51	0.12	0.00	0.53	0.00	0.47	0.00	0.00	0.408
	H_O	0.00	0.07	0.00	0.00	0.00	0.22	0.00	0.47	0.09	0.00	0.13	0.00	0.11	0.00	0.31	0.00	0.00	
Omy_128923-433	H_E	0.17	0.00	0.09	0.00	0.00	0.04	0.12	-	0.23	0.07	-	0.29	0.00	0.52	0.31	0.35	0.48	0.417
	H_O	0.18	0.00	0.10	0.00	0.00	0.04	0.04	-	0.26	0.07	-	0.20	0.00	0.60	0.38	0.43	0.57	
Omy_128996-481	H_E	0.51	0.51	0.51	0.44	0.39	0.51	0.22	0.19	0.31	0.25	0.28	0.42	0.29	0.00	0.23	0.45	0.48	0.142
	H_O	0.70	0.00	0.52	0.63	0.38	0.39	0.25	0.13	0.33	0.29	0.25	0.31	0.33	0.00	0.25	0.48	0.39	
Omy_129170-794	H_E	0.00	0.07	0.04	0.00	0.12	0.20	0.19	0.50	0.48	0.48	0.40	0.35	0.47	0.00	0.00	0.32	0.00	0.268
	H_O	0.00	0.07	0.04	0.00	0.13	0.13	0.21	0.57	0.54	0.57	0.38	0.19	0.44	0.00	0.00	0.30	0.00	
Omy_129870-756	H_E	0.46	0.07	0.49	0.31	0.39	0.37	0.50	0.50	0.37	0.43	0.50	0.51	0.29	0.29	0.39	0.37	0.46	0.137

Table 1.4 Continued

Assay Name		Willamette River, Wiley Creek	Klamath River, Buckboard Creek	Klamath River, Kelsey Creek	Goose Lake, Bauers Creek	McCloud River, Butcherknife Creek	Sacramento River, Battle Creek	Eel River, Middle Fork (summer)	Russian River	Scott Creek	San Lorenzo River	Carmel River	Eagle Lake Strain, A_RH	San Lorenzo Strain, BCH	Kamloops Strain, HCH	Mount Whitney Strain, FH	Lake Taupo, New Zealand	Río Santa Cruz, Argentina	
	N	23	15	23	16	16	24	24	47	46	15	24	16	9	15	16	23	24	F_{ST}
	H_O	0.35	0.07	0.68	0.38	0.38	0.39	0.67	0.55	0.44	0.47	0.63	0.50	0.33	0.20	0.38	0.39	0.61	
Omy_130295-98	H_E	0.46	0.50	0.47	0.18	0.42	0.45	0.38	0.12	0.22	0.07	0.00	0.18	0.00	0.48	0.27	0.45	0.50	0.629
	H_O	0.17	0.53	0.52	0.19	0.56	0.48	0.42	0.13	0.16	0.07	0.00	0.19	0.00	0.33	0.31	0.57	0.29	
Omy_130524-160	H_E	0.49	0.20	0.49	0.44	0.18	0.45	0.51	0.47	0.50	0.51	0.44	0.42	0.52	0.48	0.39	0.00	0.00	0.110
	H_O	0.78	0.21	0.55	0.38	0.19	0.42	0.54	0.36	0.43	0.47	0.46	0.31	0.44	0.47	0.38	0.00	0.00	
Omy_130720-100	H_E	0.00	0.46	0.25	0.06	0.18	0.48	0.50	0.49	0.51	0.43	0.44	0.35	0.46	0.19	0.48	0.50	0.49	0.165
	H_O	0.00	0.53	0.29	0.06	0.19	0.39	0.58	0.38	0.42	0.20	0.38	0.44	0.63	0.07	0.50	0.39	0.46	
Omy_131460-646	H_E	0.04	0.07	0.04	0.00	0.00	0.16	0.51	0.50	0.48	0.50	0.45	0.31	0.53	0.24	0.06	0.41	0.35	0.228
	H_O	0.04	0.07	0.04	0.00	0.00	0.17	0.50	0.39	0.54	0.53	0.42	0.25	0.33	0.27	0.06	0.48	0.43	
Omy_131965-120	H_E	0.20	0.33	0.47	0.23	0.00	0.50	0.49	0.51	0.48	0.46	0.31	0.50	0.52	0.29	0.51	0.50	0.36	0.047
	H_O	0.22	0.40	0.50	0.25	0.00	0.52	0.54	0.60	0.33	0.27	0.38	0.44	0.22	0.20	0.50	0.48	0.21	

Table 1.5: Gene similarity search results for the loci containing the 139 SNPs design in this study and 13 from other groups. Only E-value $\leq$ 1e-20 were consider. Location abreviations refer to UTR: Untranslated region; CDS: Coding DNA sequnce. Abreviations inside parenthesis indicate the aminoacid for the codon where the SNP is located. (aminoacid): synonymous; (aminoacid/aminoacid): non-synonymous. Bold assay names indicates the 95 SNP loci selected for the panel. *a*: Abadía-Cardoso *et al.* 2011; *b*: Aguilar & Garza 2008; *c*: Campbell *et al.* (2009); *d*: CRITFC - N. Campbell unpublished; *e*: WSU - J. DeKoning unpublished.

Assay Name	Species	Similar to:	Location (aminoacid)
OMGH1PROM1-SNP1[a]	*Oncorhynchus mykiss*	Growth hormone 1 gene, promoter region	3'UTR
Omy_95318-147[a]		n/a	Exon
Omy_95442-108[a]		n/a	Exon
Omy_95489-423[a]	*Salmo salar*	Phenazine biosynthesis-like domain-containing protein 2	Intron
Omy_96158-277[a]	*Pagrus major*	RAP2B-like protein	CDS (Pro)
Omy_96222-125[a]	*Solea senegalensis*	Elongation factor 1 alpha isoform 42Sp50	Intron
Omy_96529-231[a]	*Oncorhynchus mykiss*	Dolichol-phosphate (beta-D) mannosyltransferase 2 (dpm2)	3'UTR
Omy_96899-148[a]	*Salmo salar*	Vaccinia related kinase 3 (vrk3)	CDS (Leu/Phe)
Omy_97077-73[a]	*Salmo salar*	BTG1	3'UTR
Omy_97660-230[a]	*Salmo salar*	Single-strand selective monofunctional uracil (Smug1)	3'UTR
Omy_97865-196[a]	*Salmo salar*	n/a	3'UTR
Omy_97954-618[a]	*Oncorhynchus masou*	Retinol-binding protein 2 (RBP2)	CDS (Asn)
Omy_98188-405[a]	*Salmo salar*	Vacuolar protein sorting 72 homolog(vps72)	3'UTR
Omy_98409-549[a]	*Salmo salar*	Fizzy-related protein homolog	Intron
Omy_98683-165[a]	*Salmo salar*	Chymotrypsin-like (ctrl)	Intron
Omy_99300-202[a]		n/a	Exon
Omy_100771-63[a]		n/a	Exon
Omy_100974-386[a]		n/a	Exon
Omy_101119-554[a]		n/a	Exon
Omy_101341-188[a]	*Salmo salar*	Selenocysteine-associated protein 1 (pseudogene)	Intron
Omy_101554-306[a]	*Salmo salar*	NMDA receptor-regulated protein 1	3'UTR
Omy_101704-329[a]	*Salmo salar*	Histidyl-tRNA synthetase (pseudogene)	Intron
Omy_101770-410[a]	*Esox lucius*	Cartilage-associated protein precursor	Intron
Omy_101832-195[a]		n/a	n/a
Omy_101993-189[a]	*Equus caballus*	Syndecan-4-like	Intron
Omy_102213-204[a]	*Oncorhynchus mykiss*	Leukocyte cell-derived chemotaxin 2 precursor	Intron
Omy_102420-634[a]	*Salmo salar*	Cathepsin K precursor	Intron
Omy_102457-423[a]	*Oncorhynchus mykiss*	Heat shock protein 70a (hsp70a)	Intron
Omy_102505-102[a]		n/a	Exon
Omy_102510-682[a]		n/a	Intron
Omy_102867-443[a]		n/a	Intron
Omy_103350-395[a]	*Epinephelus coioides*	Alpha-1,3-galactosyltransferase-like protein	Exon
Omy_103577-379[a]	*Salmo salar*	NADPH–cytochrome P450 reductase	Intron

Table 1.5 Continued

Assay Name	Species	Similar to:	Location (aminoacid)
Omy_103705-558[a]		n/a	Exon
Omy_103713-53[a]		n/a	Exon
Omy_104519-624[a]	*Salmo salar*	Nucleolar transcription factor 1 (ubf1)	Intron
Omy_104569-114[a]		n/a	Exon
Omy_105075-162[a]		n/a	Exon
Omy_105105-448[a]	*Oncorhynchus mykiss*	VHSV-induced protein	Intron
Omy_105115-367[a]		n/a	Exon
Omy_105235-713[a]	*Salmo salar*	Growth arrest and DNA-damage-inducible protein GADD45 alpha	Intron
Omy_105385-406[a]		n/a	Exon
Omy_105386-347[a]		n/a	Exon
Omy_105401-363[a]		n/a	Exon
Omy_105407-74[a]	*Salmo salar*	60S ribosomal protein L36a	3'UTR
Omy_105714-265[a]	*Oncorhynchus mykiss*	FHA-HIT	CDS (Ser/Pro)
Omy_105897-101[a]	*Salmo salar*	Uridine 5-monophosphate synthase (pyr5)	3'UTR
Omy_106172-332[a]		n/a	Exon
Omy_106313-445[a]	*Salmo salar*	Wilms tumor 1 associated protein-like	Intron
Omy_106560-58[a]	*Takifugu rubripes*	Zinc finger SWIM domain-containing protein 6-like	CDS (Cys)
Omy_106747-707[a]	*Salmo salar*	CJ088 protein (cj088)	CDS (Thr)
Omy_107031-704[a]	*Oncorhynchus mykiss*	Cold inducible RNA binding protein (cirbp)	3'UTR
Omy_107074-217[a]		n/a	Exon
Omy_107285-69[a]	*Salmo salar*	Placental protein 25 (pp25)	Intron
Omy_107336-170[a]	*Salmo salar*	NUDT6	CDS (Pro/Arg)
Omy_107607-137[a]		n/a	Exon
Omy_107786-314[a]		n/a	Exon
Omy_107786-584[a]		n/a	Intron
Omy_107806-34[a]	*Salmo salar*	YIPF4	Intron
Omy_108007-193[a]		n/a	Exon
Omy_108735-311[a]		n/a	Exon
Omy_108820-85[a]		n/a	Exon
Omy_109243-222[a]	*Salmo salar*	Kunitz-type protease inhibitor 2 (spit2)	3'UTR
Omy_109390-341[a]	*Salmo salar*	Serine/threonine-protein kinase SRPK1 (pseudogene)	Exon
Omy_109525-403[a]	*Oncorhynchus mykiss*	Prostaglandin-endoperoxide synthase 2b (ptgs2b)	3'UTR
Omy_109651-445[a]	*Oncorhynchus mykiss*	Secreted phosphoprotein 24 (spp2)	3'UTR
Omy_109693-461[a]	*Salmo salar*	Nuclear transcription factor Y, gamma (nfyc)	3'UTR
Omy_109874-148[a]		n/a	Exon
Omy_109894-185[a]		n/a	Intron
Omy_109944-74[a]		n/a	Exon
Omy_110064-419[a]	*Oncorhynchus mykiss*	Lipopolysaccharide-induced TNF factor (litaf)	Intron
Omy_110078-294[a]		n/a	Exon

Table 1.5 Continued

Assay Name	Species	Similar to:	Location (aminoacid)
Omy_110201-359[a]	*Oreochromis niloticus*	Autophagy-related protein 101-like	Intron
Omy_110362-585[a]	*Salmo salar*	ATP synthase subunit alpha, mitochondrial precursor	CDS (Leu)
Omy_110571-386[a]	*Oncorhynchus mykiss*	Epithelial-cadherin	3'UTR
Omy_110689-148[a]		n/a	Exon
Omy_111005-159[a]	*Salmo salar*	RAB2A, member RAS oncogene family (rab2a)	3'UTR
Omy_111084-526[a]		n/a	Exon
Omy_111383-51[a]	*Oncorhynchus mykiss*	V-mos Moloney murine sarcoma viral oncogene-like protein	3'UTR
Omy_111666-301[a]	*Salmo salar*	Tetraspanin-16 (tsn16)	Intron
Omy_111681-432[a]	*Oncorhynchus mykiss*	NK2 homeobox 1b (nkx2.1b)	CDS (Pro)
Omy_112208-328[a]	*Oncorhynchus mykiss*	Lactate dehydrogenase B	Intron
Omy_112301-202[a]		n/a	Intron
Omy_112820-82[a]		n/a	Exon
Omy_112876-45[a]	*Oncorhynchus mykiss*	TPT1 gene for tumor protein, translationally-controlled 1	Intron
Omy_113109-205[a]	*Oncorhynchus mykiss*	Josephin-2 (jos2)	3'UTR
Omy_113128-73[a]		n/a	Exon
Omy_113242-163[a]		n/a	Exon
Omy_113490-159[a]		n/a	Exon
Omy_114315-438[a]	*Salmo salar*	Tumor protein D54	3'UTR
Omy_114448-87[a]		n/a	Exon
Omy_114587-480[a]		n/a	Intron
Omy_114976-223[a]	*Salmo salar*	Canopy homolog 2 precursor	3'UTR
Omy_115987-812[a]		n/a	Exon
Omy_116104-229[a]		n/a	Intron
Omy_116362-467[a]	*Salmo salar*	Adhesion regulating molecule 1 ADRM1	CDS (Ile/Met)
Omy_116733-349[a]		n/a	Intron
Omy_116938-264[a]		n/a	Exon
Omy_117242-419[a]		n/a	Exon
Omy_117259-96[a]		n/a	Exon
Omy_117286-374[a]		n/a	Exon
Omy_117370-400[a]	*Danio rerio*	EST-domain transcription factor PEA3	Intron
Omy_117432-190[a]	*Takifugu rubripes*	UDP-glucuronosyltransferase 2B1-like	CDS (Ile)
Omy_117540-259[a]	*Salmo salar*	Cyclin I (ccni)	3'UTR
Omy_117549-316[a]		n/a	Exon
Omy_117743-127[a]		n/a	Exon
Omy_117815-81[a]	*Salmo salar*	Glutaredoxin (thioltransferase) (glrx)	3'UTR
Omy_118175-396[a]		n/a	Exon
Omy_118205-116[a]	*Salmo salar*	Sjoegren syndrome/scleroderma autoantigen 1 homolog	CDS (Leu)
Omy_118654-91[a]		n/a	Exon
Omy_118938-341[a]		n/a	Exon

Table 1.5 Continued

Assay Name	Species	Similar to:	Location (aminoacid)
Omy_119108-357[a]	*Salmo salar*	CK046 protein (ck046)	CDS (Leu)
Omy_119892-365[a]	*Salmo salar*	DnaJ homolog subfamily C member 1 (dnjc1)	Intron
Omy_120255-332[a]	*Salmo salar*	C6orf64	Intron
Omy_120950-569[a]		n/a	Intron
Omy_121006-131[a]	*Salmo salar*	Immediate early response gene 5 protein	CDS (Pro/Thr)
Omy_121713-115[a]		n/a	Exon
Omy_123044-128[a]		n/a	Exon
Omy_123048-119[a]	*Salmo salar*	Ribosomal protein S26	Intron
Omy_123921-144[a]		n/a	Intron
Omy_124774-530[a]	*Salmo salar*	cAMP-responsive element-binding protein 3-like protein	3'UTR
Omy_125998-61[a]		n/a	Exon
Omy_126160-242[a]		n/a	Exon
Omy_127236-583[a]		n/a	Exon
Omy_127510-920[a]	*Danio rerio*	si:dkey-21n12.1	Intron
Omy_127645-308[a]		n/a	Exon
Omy_127760-385[a]	*Salmo salar*	Zinc finger protein 503 (zn503), mRNA	Intron
Omy_128302-430[a]	*Oncorhynchus mykiss*	Ribosomal protein L20, mitochondrial precursor	Intron
Omy_128693-755[a]	*Salmo salar*	Stathmin-like 4 (stmn4)	Intron
Omy_128851-273[a]	*Oncorhynchus mykiss*	Differentially regulated trout protein 1	Intron
Omy_128923-433[a]		n/a	Exon
Omy_128996-481[a]		n/a	Exon
Omy_129170-794[a]		n/a	Exon
Omy_129870-756[a]		n/a	Exon
Omy_130295-98[a]		n/a	Exon
Omy_130524-160[a]		n/a	Exon
Omy_130720-100[a]		n/a	Exon
Omy_131460-646[a]	*Salmo salar*	Neural cell adhesion molecule L1-like protein precursor	Intron
Omy_131965-120[a]		n/a	Exon
Omy_PEPA-INT6[b]	*Oncorhynchus mykiss*	Nonspecific dipeptidase gene	3'UTR
OmyAldA[b]	*Oncorhynchus mykiss*	Aldolase A gene, intron 1	Intron
ONMYCRBF_1-SNP1[b]	*Oncorhynchus mykiss*	Carbonyl reductase/20beta-hydroxysteroid dehydrogenase A	3'UTR
Omy_arp-630[c]	*Oncorhynchus mykiss*	Acidic ribosomal phosphoprotein	3'UTR
Omy_aspAT-123[c]	*Oncorhynchus mykiss*	Aspartate aminotransferase	3'UTR
Omy_gh-475[c]	*Oncorhynchus mykiss*	Growth hormone 1	3'UTR
Omy_nramp-146[c]	*Oncorhynchus mykiss*	Natural resistance-associated macrophage protein-alpha (nramp-a)	3'UTR
Omy_Ogo4-212[c]	*Oncorhynchus gorbuscha*	Microsatellite locus Ogo4, (gt)55	n/a
OmyCOX1-221[c]	*Oncorhynchus mykiss*	mRNA for cyclooxygenase-1 (cox-1)	3'UTR
Omy_mapK3-103[d]	No sequence available		
Omy_g12-82[e]	No sequence available		

Table 1.5 Continued

Assay Name	Species	Similar to:	Location (aminoacid)
Omy_gsdf-291[e]	No sequence available		
Omy_mcsf-371[e]	No sequence available		

1.4 Discussion

I report the discovery and development of assays for 139 novel single nucleotide polymorphisms in the species *O. mykiss*, steelhead/rainbow trout, through sequence analysis of 236 ESTs with a total consensus length of 120KB. I demonstrate how ESTs from existing public databases and directed Sanger-sequencing of PCR products can be used to identify large numbers of SNPs in non-model organisms. In species and populations with large effective sizes, such sequencing from existing genomic information uncovers sufficient polymorphism that a preliminary screen of loci for potential polymorphism, using methods such as single strand conformation polymorphism or high resolution melt analyses, can be avoided, since nearly every locus will contain some variants.

The 139 SNP loci described here are broadly polymorphic in the species and should prove useful for a variety of applications, including phylogeography, genetic stock identification, individual identification, behavioral ecology and pedigree reconstruction. The availability of large numbers of SNPs known to be polymorphic in populations of steelhead and rainbow trout will allow the implementation of intergenerational genetic tagging through large-scale parentage inference, since this requires only about 100 SNP loci for sufficiently low tag recovery error rates (Garza & Anderson 2007). Such parentage-based tagging (PBT) will allow an unprecedented level of monitoring and evaluation of natural and hatchery/aquaculture populations, including estimation of variance in reproductive success, migration rates, effective population sizes, life-stage

specific mortality rates, and other population parameters. PBT is based on the principle that genotyping fish from the parental generation, either in a hatchery, an aquaculture operation or a natural population, provides intergenerational genetic tags for their progeny that can be retrieved through large-scale parentage inference (Anderson & Garza 2006; Garza & Anderson 2007). Such pedigree reconstruction is greatly facilitated by the low genotyping error/mutation rates of SNP loci. In addition, as more SNP loci are described and more assays become available for the species, it will be possible to construct second-generation genetic linkage maps and high-density SNP genotyping microarrays become available the pedigrees resulting from PBT will enable detailed understanding of the genetic architecture of phenotypic traits in the species. Because of its importance in recreational fisheries and in aquaculture, as well as the ESA protection of many populations, the species *O. mykiss* is among the most economically significant fishes in the world, and an increased understanding of its phenotypic variation is of great value.

During the past decade, microsatellite markers have dominated population genetic work in salmonids, due to their high variability and conservation among related species (Aguilar & Garza 2006; Clemento *et al.* 2009; Pearse *et al.* 2007; Pearse *et al.* 2009). However, microsatellites have significant drawbacks, among them relatively high genotyping error/mutation rates, significant staff time necessary for data generation and allele calling, and homoplasy. Moreover, the results obtained with microsatellites in one laboratory are not directly combinable with data generated in other laboratories, even when using the same instrumentation, due to subtle differences in electrophoretic

conditions and consequent data output (Seeb *et al.* 2007). The requirement for a standardization process to be able to combine microsatellite data between laboratories adds significant time and expense to collaborative projects.

Conversely, data obtained from SNP loci are easily portable and combinable between labs, as long as the same primer/probe sequences and/or reporting conventions are used. This will allow large multilateral databases to be developed for applications in fishery management, ecological investigation, and aquaculture/hatchery broodstock management using both standard (*e.g.* Seeb *et al.* 2007) and pedigree-based approaches (Anderson & Garza 2006). Moreover, the advent of new technologies, such as nanofluidic circuitry and spotted arrays, for thermal cycling and genotyping now allows the examination of a large number of SNPs in a large number of individuals in a short time period and at relatively low cost. This provides the prospect of SNP genotyping as a routine, and very valuable, tool for monitoring and evaluation of steelhead and rainbow trout populations throughout the world.

As SNP loci are typically bi-allelic, the amount of information per locus is more limited than for most multiallelic loci, such as microsatellites or AFLPs. In the future, however, analysis of haplotypes of tightly linked SNPs may provide additional information for many questions, including in phylogeography and pedigree resolution. Since I discovered many additional polymorphic sites in these genes, it would be possible to design additional assays for many of these sites and perform haplotype analyses. More complete analyses of this sequence variability will be reported elsewhere.

The number and density of substitutions and SNPs discovered here was con-

sistent with what has been reported for other salmonids (Smith *et al.* 2005), but it is difficult to draw direct comparisons between different SNP discovery efforts, since the density of polymorphic sites uncovered depends critically on the number and phylogenetic diversity of the individuals in the ascertainment panel, the set of genes or genomic sequences interrogated for SNP discovery, and accuracy of the sequencing method employed. My ascertainment approach and stringent design criterion for SNP discovery were intended to fulfill several objectives. Included in the ascertainment panel were both representatives from populations in California where I are actively working and intend to apply the resulting markers, as well as from rainbow trout strains commonly used throughout the world for fishery stocking and/or aquaculture. By designing assays for variable sites only when all three genotypes were observed, and without regard to which individuals carried them, I selected both for markers with a higher mean MAF and markers that were more likely to be broadly useful in the species. This was intended to provide markers useful both for study and management of native steelhead populations, as well as with the millions of rainbow trout cultured for food and fisheries. However, it will also underrepresent rare variants, which could result in biases in phylogenetic and evolutionary applications of these markers. Still, it is important to point out that sets of microsatellite and other population genetic markers developed for salmonids and other non-model organisms suffer from the same biases. Therefore, applications of these SNP markers that depend upon a representative sampling of the site frequency spectrum in focal populations or lineages should ideally employ markers ascertained using diverse ascertainment populations and strategies.

My ascertainment panel included fish from three coastal steelhead populations from several closely-related lineages, a highly divergent population of redband trout and several rainbow trout strains domesticated from distinct lineages. This diverse ascertainment panel was intended to reduce ascertainment bias in populations in the southern part of the North American range.

1.5 References

Aguilar A, Garza JC (2006) A comparison of variability and population structure for major histocompatibility complex and microsatellite loci in California coastal steelhead (*Oncorhynchus mykiss* Walbaum). Molecular Ecology 15, 923-937.

Aguilar A, Garza JC (2008) Isolation of 15 single nucleotide polymorphisms from coastal steelhead, *Oncorhynchus mykiss* (Salmonidae). Molecular Ecology Resources 8, 659-662.

Anderson EC, Garza JC (2006) The power of single-nucleotide polymorphisms for large-scale parentage inference. Genetics 172, 2567-2582.

Beacham TD, Pollard S, Le KD (2000) Microsatellite DNA population structure and stock identification of steelhead trout (*Oncorhynchus mykiss*) in the Nass and Skeena Rivers in northern British Columbia. Marine Biotechnology 2, 587-600.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la gntique des populations. Laboratoire Gnome, Populations, Interactions, CNRS UMR 5000, Universit de Montpellier II, Montpellier (France).

Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. Trends in Ecology & Evolution 18, 249-256.

Burgner RL, Light JT, Margolis L, *et al.* (1992) Distribution and origins of steelhead trout *Oncorhynchus mykiss* in offshore waters of the north Pacific Ocean. International North Pacific Fisheries Commission Bulletin, 1-92.

Busack CA, Gall GAE (1980) Ancestry of artificially propagated California rainbow trout strains. California Fish and Game 66, 17-24.

Busby PJ, Wainwright TC, Bryant GJ, *et al.* (1996) Status review of west coast steelhead from Washington, Idaho, Oregon, and California. NOAA Technical Memorandum NMFS-NWFSC-27.

Campbell NR, Overturf K, Narum SR (2009) Characterization of 22 novel single nucleotide polymorphism markers in steelhead and rainbow trout. Molecular Ecology Resources 9, 318-322.

Castaño-Sánchez C, Smith TPL, Wiedmann RT, *et al.* (2009) Single nucleotide polymorphism discovery in rainbow trout by deep sequencing of a reduced representation library. BMC Genomics 10, 559.

Clark AG, Hubisz MJ, Bustamante CD, Williamson SH, Nielsen R (2005) Ascertainment bias in studies of human genome-wide polymorphism. Genome Research 15, 1496-1502.

Clemento AJ, Anderson EC, Boughton D, Girman D, Garza JC (2009) Population genetic structure and ancestry of *Oncorhynchus mykiss* populations above and below dams in south-central California. Conservation Genetics 10, 1321-1336.

Fausch KD (2007) Introduction, establishment and effects of non-native salmonids: considering the risk of rainbow trout invasion in the United Kingdom. Journal of Fish Biology 71, 1-32.

Garza JC, Anderson EC (2007) Large scale parentage inference as an alternative to coded-wire tags for salmon fishery management in Pacific salmon commission GSI workshops: logistics workgroup final report and recommendations, p. 55.

Garza JC, Pearse DE (2008) Population genetic structure of *Oncorhynchus mykiss* in the California Central Valley. Report to the California Department of Fish and Game p. 54.

Gene Index. http://compbio.dfci.harvard.edu Dana-Farber Cancer Institute and Harvard School of Public Health.

Martínez A, Garza JC, Pearse DE (2011) A microsatellite genome screen identifies chromosomal regions under differential selection in steelhead and rainbow trout (*Oncorhynchus mykiss*). Transactions of the American Fisheries Society 140, 829-842.

McCarthy JJ, Hilfiker R (2000) The use of single-nucleotide polymorphism maps in pharmacogenomics. Nature Biotechnology 18, 505-508.

McCusker MR, Parkinson E, Taylor EB (2000) Mitochondrial variation in rainbow trout *Oncorhynchus mykiss* across its native range: testing biogeographical hypotheses and their relevance to conservation. Molecular Ecology 9, 2089-2108.

Narum SR, Banks M, Beacham TD, *et al.* (2008) Differentiating salmon populations at broad and fine geographical scales with microsatellites and single nucleotide polymorphisms. Molecular Ecology 17, 3464-3477.

Narum SR, Powell MS, Talbot AJ (2004) A distinctive microsatellite locus that differentiates ocean-type from stream-type Chinook salmon in the interior Columbia River basin. Transactions of the American Fisheries Society 133, 1051-1055.

NOAA-National Oceanographic and Atmospheric Administration (2006) Endangered and Threatened species: final listing determinations for 10 Distinct Population Segments of west coast steelhead. US Federal Register 71, 833-862.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89, 583-590.

Pascual M, Bentzen P, Rossi CR, *et al.* (2001) First documented case of anadromy in a population of introduced rainbow trout in Patagonia, Argentina. Transactions of the American Fisheries Society 130, 53-67.

Pearse DE, Donohoe CJ, Garza JC (2007) Population genetics of steelhead (*Oncorhynchus mykiss*) in the Klamath River. Environmental Biology of Fishes 80, 377-387.

Pearse DE, Hayes SA, Bond MH, *et al.* (2009) Over the falls? Rapid evolution of ecotypic differentiation in steelhead/rainbow trout (*Oncorhynchus mykiss*). Journal of Heredity 100, 515-525.

Rexroad III CE, Palti Y, Gahr SA, Vallejo RL (2008) A second generation genetic map for rainbow trout (*Oncorhynchus mykiss*). BMC Genetics 9.

Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources 8, 103-106.

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Bioinformatics methods and protocols: methods in molecular biology (eds. Krawetz S, Misener S), pp. 365-386. Humana Press, Totowa, NJ.

Sachidanandam R, Weissman D, Schmidt SC, *et al.* (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409, 928-933.

Seeb LW, Antonovich A, Banks MA, *et al.* (2007) Development of a standardized DNA database for Chinook salmon. Fisheries 32, 540-552.

Shapovalov LEO, Taft AC (1954) The life histories of the steelhead rainbow trout (*Salmo gairdneri gairdneri*) and silver salmon (*Oncorhynchus kisutch*) with special reference to Waddell Creek, California, and recommendations regarding their management. California Department of Fish and Game Fish Bulletin 98, 5-375.

Smith CT, Elfstrom CM, Seeb LW, Seeb JE (2005) Use of sequence data from rainbow trout and Atlantic salmon for SNP detection in Pacific salmon. Molecular Ecology 14, 4193-4203.

Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other types of molecular markers and their use in animal genetics. Genetics Selection Evolution 34,

275-305.

Wang DG, Fan JB, Siao CJ, *et al.* (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280, 1077-1082.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38, 1358-1370.

Young WP, Ostberg CO, Keim P, Thorgaard GH (2001) Genetic characterization of hybridization and introgression between anadromous rainbow trout (*Oncorhynchus mykiss irideus*) and coastal cutthroat trout (*O. clarki clarki*). Molecular Ecology 10, 921-930.

Chapter 2

Large-scale parentage analysis reveals reproductive patterns and heritability of spawn timing in a hatchery population of steelhead (*Oncorhynchus mykiss*)

Abstract

Understanding life history traits is an important first step in formulating effective conservation and management strategies. The use of artificial propagation and supplementation as such a strategy can have numerous effects on the supplemented natural populations and minimizing life history divergence is crucial in minimizing these effects. Here, I use SNP genotypes for large-scale parentage analysis and pedigree reconstruction in a hatchery population of steelhead, the anadromous form of rainbow trout. Nearly complete sampling of the broodstock for several consecutive years in two hatchery programs allowed inference about multiple aspects of life history. Reconstruction of cohort age distribution revealed a strong component of fish that spawn at two years of age, in contrast to program goals and distinct from naturally spawning steelhead in the region, which raises a significant conservation concern. The first estimates of variance in family size for steelhead in this region can be used to calculate effective population size and probabilities of inbreeding and estimation of iteroparity rate, indicates that it is reduced by hatchery production. Finally, correlations between family members in the day of spawning revealed for the first time a strongly heritable component to this important life history trait in steelhead and demonstrated the potential for selection to alter life history traits rapidly in response to changes in environmental conditions. Taken together, these results demonstrate the extraordinary promise of SNP-based pedigree reconstruction for providing biological inference in high-fecundity organisms that is not easily achievable with traditional physical tags.

2.1 Introduction

Understanding a population's life history patterns, and the environmental and biological factors that affect them, is a first step in the formulation of effective conservation and management actions. For example, lack of knowledge regarding when and where an animal disperses could lead to incorrect habitat conservation priorities. In addition, life history traits such as reproductive strategy and migratory behavior shape the direction of evolution and responses to environmental change (Hansen *et al.* 2012). Such responses are due to selection on the heritable variation in life history traits, yet the components of observed trait values that are due to genetic variation and phenotypic plasticity are even less frequently known. However, even observable components of life history traits can be hard to measure in species with complex life cycles, high fecundity and/or a high degree of dispersal or migration (*e.g.* anadromous fish, many birds, etc.). As many populations of animal and plant species decline in numbers and in geographic distribution, a clear understanding of their life histories and reproductive biology is critical to prevent further extirpations and extinction.

Artificial propagation, followed by supplementation of natural populations, is a widely employed method for addressing population declines (Champagnon *et al.* 2012). Artificially propagated populations may face a range of problems, including those resulting from what is generally termed domestication selection, as well as inbreeding depression, increased disease susceptibility, etc. (Bryant & Reed 1999; Swinnerton *et al.* 2004). This gives rise to significant concerns about genetic effects, as natural populations

often experience maladaptation and reduction in fitness due to introgression by stocked individuals (Utter 1998; Araki *et al.* 2007; Frankham 2008; Williams & Hoffman 2009). A potential strategy to minimize negative fitness effects is to avoid artificial selection as much as possible in the captive population (Frankham *et al.* 2002; Baskett & Waples 2013). This requires explicit knowledge about both life history traits in the species and the extent to which the propagated population differs from the natural one.

Fishes in the family Salmonidae are perhaps the world's most commonly propagated organisms for which the goal of captive production is supplementation of natural populations. The scope of captive production of salmonids is vast, with at least 800 hatcheries releasing fish into tributaries of the North Pacific Ocean alone (Augerot 2005). While salmonids are among the more intensively studied animal species, there are still substantial gaps in our knowledge of basic life history, particularly on the periphery of the native geographic distribution of Pacific salmonids (Quinn 2004). Because of the extraordinary amount of phylogeographic structure and local adaptation in this group, values of life history traits often differ even between geographically proximate populations and inference drawn in one population can not necessarily be extrapolated to another. For example, the timing of reproduction of Coho salmon (*Oncorhynchus kisutch*) is clinal at the southern end of their range, with the mean date of spawning varying by more than two months over less than 500 Km (Weitkamp *et al.* 1995).

Populations of salmon and steelhead on the west coast of North America have declined dramatically over the past century and many populations are now protected under the United States Endangered Species Act (ESA; NOAA 2006). Supplementa-

tion with billions of hatchery-produced juvenile fish has not only failed to stabilize or increase many salmonid populations, but may have actually contributed to their decline (Augerot 2005). Mismatches between mean values of traits in supplemented populations and environmental conditions to which the natural population was initially adapted can cause dramatic decreases in fitness (Utter 1998; Frankham 2008; Palkovacs *et al.* 2012). For example, the selection of spawners at a hatchery may not replicate patterns of relative reproductive success in nature (Hoffnagle *et al.* 2008), where intraspecific competition and other behavioral traits play an important role (Fleming 1998), and fish that would have low fitness in natural spawning may produce many progeny. Such hatchery management practices may contribute to a reduction in genetic variability and fitness of the population (Araki *et al.* 2008) in as little as a single generation of captive breeding (Christie *et al.* 2012). Without a detailed understanding of their life history and reproductive biology guiding hatchery practices, such supplementation will almost inevitably have negative fitness consequences on the associated natural populations.

Amongst salmonids, *O. mykiss* has perhaps the most variability in life history (Shapovalov & Taft 1954; Busby *et al.* 1996). Two major ecotypes of *O. mykiss* can be distinguished: the anadromous type called "steelhead" and the nonanadromous resident type called rainbow or redband trout. After hatching, steelhead spend one to seven years in freshwater and one or more years in the ocean before returning, usually to their natal stream, to spawn. In addition to substantial variation in timing of freshwater entry and associated reproductive maturity and spawn timing, steelhead may also be iteroparous, spawning in more than one year (Busby *et al.* 1996). This complex life history makes

both understanding their biology and effectively managing their populations a significant challenge.

Numerous studies on life history variation, survival, and migration of *O. mykiss* have been undertaken (Busby *et al.* 1996), including work on the inheritance of life history traits (Thrower *et al.* 2004; Nichols *et al.* 2008; Martínez *et al.* 2011; Hecht *et al.* 2012; Miller *et al.* 2012). Spawn timing has been shown to be heritable in several salmonid species (Hendry & Day 2005; Carlson & Seamons 2008). However, due to the difficulties noted above, most estimates of heritability of spawn timing have been in entirely captive families (Siitonen & Gall 1989; Su *et al.* 1997; Wilson *et al.* 2003; Haidle *et al.* 2008; Colihueque *et al.* 2010) and very few have been in free-living salmonids (Smoker *et al.* 1998; Quinn *et al.* 2000; Dickerson *et al.* 2005). Studying inheritance in anadromous populations is challenging because of the difficulty of tracking families through their ocean migrations from one generation to the next and heritability of spawn timing has not been studied in steelhead.

Early studies on anadromous salmonids employed a variety of external marks (*e.g.* fin clips, maxillary clips, etc.) and later coded wire tags (CWTs; Hankin *et al.* 2005) to distinguish families, but these methods require considerable labor and typically allow identification of only a small number of families. They also often require sacrificing the fish for individual identification (Cooke *et al.* 2004; Hankin *et al.* 2005), which is poor practice in iteroparous species. Improvements in molecular markers and statistical analyses have provided a cost-effective alternative to traditional tagging methods, by using genotype data to identify previously sampled individuals and to identify their

kin by reconstructing pedigrees (*e.g.* Palsbøll 1999; Pearse *et al.* 2001; Blouin 2003; Garrigue *et al.* 2004). The application of parentage analysis as a tagging method is particularly powerful, as it allows direct identification of a genotyped individual's progeny (and parents), without having to "tag" any fish that will not be resampled. Such approaches have been very useful in understanding biological patterns at the population and individual levels (Avise *et al.* 2002; Planes *et al.* 2009; Hudy *et al.* 2010).

Recently, single nucleotide polymorphisms (SNPs) have emerged as reliable, cost-effective genetic markers that are easily developed by leveraging recent advances in sequencing technology and genomic resources (*e.g.* Abadía-Cardoso *et al.* 2011; Clemento *et al.* 2011). Although SNPs were initially granted dim prospects for relationship inference in molecular ecology (Glaubitz *et al.* 2003), Anderson and Garza (2006) demonstrated that a relatively small number of SNPs (< 100) would allow accurate parentage studies larger than any that had been previously attempted. The coincidence of that work with the advent of novel genotyping platforms that permit the rapid genotyping of thousands of individuals at many loci has now set the stage for SNPs to be the marker of choice for large-scale parentage studies and for genetic tagging of migratory species.

Here I examine whether a pedigree-based intergenerational genetic tagging protocol can provide information comparable to that provided by physical tagging methods and use it to elucidate reproductive patterns in ESA-listed steelhead from a supplementation program in the Russian River, CA, USA. Specifically, I determine if I can assign most individuals that return from the ocean to pairs of parents that were spawned on

the same day, but without cross information recorded. I then use the resulting parent-offspring trios to estimate the age distribution and variation in family size (*i.e.* number of siblings) amongst reproducing fish. I ask whether fish of different ages spawn on significantly different dates and use the pedigrees to provide the first estimates of the heritability of date of spawning in steelhead. A matching samples analysis allows us to estimate the number of fish that are spawned multiple times within a single season and the number that return and reproduce in multiple seasons, and ask whether these estimates are consistent with program goals and what has been observed in other populations of the species.

These hatchery populations use local fish as broodstock, provide substantial numbers of spawners in natural areas and are not genetically differentiated from the natural populations in the Russian River (Deiner *et al.* 2007). As such, elucidation of life history patterns in these hatchery steelhead populations allows us to examine whether they may be negatively influencing the associated natural populations. I demonstrate how the use of pedigree-based genetic tagging provides a powerful means of understanding many basic biological traits in relatively high fecundity species with significant conservation concerns.

2.2 Methods

2.2.1 *Study system*

The Russian River drains into the Pacific Ocean approximately 100 km north of San Francisco Bay, USA (Figure 2.1). It supports populations of Chinook salmon (*O. tshawytscha*), coho salmon and steelhead. Construction of two large dams in 1959 and 1982 blocked access to spawning and rearing habitat in the basin and two hatchery programs were established to mitigate these losses: Warm Springs (Don Clausen) Hatchery (WSH) at the base of Warm Springs Dam, on the mid-basin Dry Creek tributary, and Coyote Valley Fish Facility (CVFF) located below Coyote Valley Dam, near the headwaters (Figure 2.1).

Adult steelhead enter the Russian River to spawn from December to April. Broodstock were chosen from amongst all returning adults without regard to phenotypic characteristics, except that fish below ∼50 cm in length, presumably resident and age-two fish, were excluded. Broodstock at CVFF and WSH were mated with an approximate male to female ratio of two to one, and three to one, respectively. Each male is supposed to be crossed with only one female, but when there are not enough males to accomplish this, a previously spawned male may be reused with a different mate. All fish were released back into the river after spawning and could potentially enter the hatchery again. Fish trapped at the two facilities were spawned separately at WSH, and all juveniles were initially incubated there. Juveniles produced from CVFF adults were then moved to CVFF for imprinting, since olfactory cues experienced in early life

are critical to proper homing in salmonids (Cooper & Scholz 1976). All fish were then released at age-one. All juvenile fish produced at WSH and CVFF were marked prior to release by adipose fin removal, allowing easy visual determination of hatchery origin. There are no other steelhead hatchery programs in the vicinity of the Russian River, so nearly all fish lacking an adipose fin should have originated in one of these two programs.

2.2.2 *Tissue collection and DNA extraction*

Small ($\approx$1cm^2) caudal fin clips were collected from 3,546 adult steelhead at WSH from 2007 to 2011 and from 1,450 adult steelhead at CVFF from 2009 to 2011. These samples are believed to represent all the individuals that were spawned during those years, except for one spawning day in 2008 at WSH when 18 males and six females were spawned but not sampled.

Tissue samples were digested with proteinase K, followed by DNA extraction with DNeasy 96 Tissue Kits (QIAGEN Inc.). Purified DNA was diluted 1:2 in ddH_2O prior to a pre-amplification PCR with primers derived from 96 real time assays to enrich the DNA fragments containing the loci of interest. PCR reagent concentrations and thermal cycling conditions are available from the authors upon request.

2.2.3 *SNP loci and genotyping*

A panel of 95 SNPs was selected from a set of 192 loci (Aguilar & Garza 2008; Campbell *et al.* 2009; Abadía-Cardoso *et al.* 2011), based on their utility for parentage inference in four steelhead populations in California (including WSH) and their ability

Figure 2.1: Location of the Warm Springs (Don Clausen) Hatchery (WSH) at the base of Warm Springs Dam, on the mid-basin Dry Creek tributary, and Coyote Valley Fish Facility (CVFF) located below Coyote Valley Dam.

to distinguish individuals from several California populations (unpublished data). In the Russian River, 18 of these SNPs have a minor allele frequency < 0.15, 34 between 0.15 and 0.3, and $43 > 0.3$ (Table 2.1). In addition, a sex identification assay consisting of a Y chromosome-linked gene probe developed by Brunelli *et al.* (2008) and an invariant autosomal gene was included in the panel to determine genetic sex of all fish.

All SNP genotyping used TaqMan assays (Applied Biosystems) on 96.96 Dynamic Genotyping Arrays with the EP1 Genotyping System (Fluidigm Corporation). Two negative (no template) controls were included in each array, and genotypes were scored using SNP Genotyping Analysis Software v3.1.1 (Fluidigm).

Table 2.1: Genotyping assays used in this study. Expected (H_E) and observed (H_O) heterozygosity, and minor allele frequency (MAF) expressed as percentage. WSU: Washington State University; CRITFC: Columbia River Inter-Tribal Fish Commission.

Assay ID	Reference	H_E	H_O	MAF
OMGH1PROM1-SNP1	Abadía-Cardoso *et al.* 2011	0.478	0.462	39.49
SH100771-63	“ ”	0.462	0.465	36.16
SH100974-386	“ ”	0.306	0.312	18.87
SH101554-306	“ ”	0.458	0.462	35.53
SH101770-410	“ ”	0.384	0.378	25.89
SH101832-195	“ ”	0.499	0.507	47.82
SH101993-189	“ ”	0.438	0.437	32.41
SH102420-634	“ ”	0.485	0.504	41.33
SH102505-102	“ ”	0.361	0.342	23.67
SH102510-682	“ ”	0.300	0.305	18.38
SH102867-443	“ ”	0.464	0.448	36.58
SH103350-395	“ ”	0.478	0.480	39.42
SH103577-379	“ ”	0.357	0.354	23.23
SH103705-558	“ ”	0.348	0.343	22.40
SH104519-624	“ ”	0.478	0.471	39.50
SH105075-162	“ ”	0.261	0.266	15.45
SH105105-448	“ ”	0.415	0.417	29.39
SH105115-367	“ ”	0.289	0.282	17.53
SH105385-406	“ ”	0.490	0.494	42.83
SH105386-347	“ ”	0.155	0.149	8.44
SH105714-265	“ ”	0.364	0.347	23.89
SH106172-332	“ ”	0.065	0.063	3.36
SH106313-445	“ ”	0.476	0.482	39.11
SH107074-217	“ ”	0.460	0.458	35.84
SH107285-69	“ ”	0.329	0.316	20.73
SH108735-311	“ ”	0.395	0.386	27.06
SH109243-222	“ ”	0.252	0.237	14.78
SH109525-403	“ ”	0.472	0.475	38.12
SH109651-445	“ ”	0.178	0.176	9.89
SH109693-461	“ ”	0.490	0.489	43.00
SH109874-148	“ ”	0.203	0.188	11.44

Table 2.1 Continued

Assay ID	Reference	H_E	H_O	MAF
SH110064-419	“ ”	0.230	0.198	13.26
SH110078-294	“ ”	0.321	0.327	20.10
SH110201-359	“ ”	0.486	0.474	41.66
SH110362-585	“ ”	0.482	0.485	40.55
SH110689-148	“ ”	0.472	0.480	38.13
SH111666-301	“ ”	0.500	0.497	48.50
SH112208-328	“ ”	0.480	0.483	40.03
SH112301-202	“ ”	0.036	0.036	1.82
SH112820-82	“ ”	0.480	0.488	39.96
SH113109-205	“ ”	0.346	0.351	22.29
SH113128-73	“ ”	0.482	0.482	40.39
SH114315-438	“ ”	0.307	0.310	18.93
SH114448-87	“ ”	0.368	0.384	24.34
SH114587-480	“ ”	0.071	0.070	3.66
SH114976-223	“ ”	0.402	0.375	27.82
SH115987-812	“ ”	0.479	0.470	39.74
SH116733-349	“ ”	0.245	0.247	14.32
SH117259-96	“ ”	0.424	0.427	30.47
SH117286-374	“ ”	0.102	0.099	5.37
SH117370-400	“ ”	0.469	0.455	37.61
SH117540-259	“ ”	0.317	0.310	19.76
SH117815-81	“ ”	0.375	0.382	25.04
SH118175-396	“ ”	0.301	0.303	18.46
SH118654-91	“ ”	0.489	0.498	42.47
SH118938-341	“ ”	0.337	0.314	21.45
SH119108-357	“ ”	0.425	0.400	30.62
SH119892-365	“ ”	0.495	0.490	45.13
SH120255-332	“ ”	0.374	0.374	24.92
SH120950-569	“ ”	0.336	0.260	21.39
SH121006-131	“ ”	0.370	0.381	24.45
SH123044-128	“ ”	0.499	0.486	48.06
SH125998-61	“ ”	0.493	0.488	44.03
SH127236-583	“ ”	0.443	0.443	33.09
SH127510-920	“ ”	0.324	0.314	20.32
SH127645-308	“ ”	0.001	0.001	0.06
SH128851-273	“ ”	0.244	0.211	14.25
SH128996-481	“ ”	0.324	0.325	20.34
SH129870-756	“ ”	0.492	0.493	43.68
SH130524-160	“ ”	0.498	0.484	46.63
SH130720-100	“ ”	0.481	0.470	40.36
SH131460-646	“ ”	0.427	0.424	30.83
SH131965-120	“ ”	0.393	0.349	26.89
SH95318-147	“ ”	0.402	0.384	27.83
SH95489-423	“ ”	0.439	0.426	32.58
SH96222-125	“ ”	0.413	0.412	29.15
SH97077-73	“ ”	0.312	0.313	19.35
SH97954-618	“ ”	0.498	0.478	46.77
SH98188-405	“ ”	0.240	0.239	13.97
SH98409-549	“ ”	0.499	0.510	48.01
SH98683-165	“ ”	0.441	0.430	32.85
SH99300-202	“ ”	0.084	0.084	4.40
Omy_AldA	Aguilar & Garza 2008	0.293	0.289	17.86
OMY_PEPA-INT6	“ ”	0.444	0.385	33.27
ONMYCRBF_1-SNP1	“ ”	0.460	0.462	35.91
*SEX_ID	Brunelli *et al.* 2008	0.451	0.687	34.37
Omy_arp-630	Campbell *et al.* 2009	0.326	0.329	20.50
Omy_aspAT-123	“ ”	0.156	0.143	8.55
Omy_COX1-221	“ ”	0.496	0.500	45.25
Omy_gh-475	“ ”	0.294	0.292	17.88

Table 2.1 Continued

Assay ID	Reference	H_E	H_O	MAF
Omy_nramp-146	“ ”	0.153	0.153	8.34
Omy_Ogo4-304	“ ”	0.331	0.318	20.96
Omy_mapK3-103	CRITFC - N. Campbell unpubl.	0.500	0.513	49.54
Omy_g12-82	WSU - J. DeKoning unpubl.	0.444	0.445	33.21
Omy_gsdf-291	“ ”	0.210	0.210	11.94
Omy_mcsf-371	“ ”	0.226	0.224	13.00

*SEX_ID Assay details (primers and probes):
Autosomal Marker F (OmyA F)
Sequence: 5’- GCC TGC TTG CAG AAG TTT TT -3’
Autosomal Marker R (OmyA R)
Sequence: 5’- CTT GAC TGT GTC CAG CTT GC -3’
Sex-linked Marker 1.4 F (OmyY1.4 F)
Sequence: 5’- CAC AAC ATG AGC TCA TGG G -3’
Sex-linked Marker 1 R (OmyY1 R)
Sequence: 5’- CGA TTA GAA AGG CCT GCT TG -3’
Autosomal Probe (OmyA probe e500)
Sequence: VIC-GAG GGG TAG TCG TTT GTT CG-MGBNFQ
Sex-linked Probe V2 (OmyY1 probe e2)
Sequence: 6FAM-CCT ACC AAG TAC AGC CCC AA-MGBNFQ

2.2.4 *Matching samples and iteroparity rate*

Samples with identical genotypes were identified to enumerate a) iteroparous individuals -those that spawned in more than one year, and b) individuals spawned multiple times within the same year. All samples with matching genotypes were re-genotyped with a second DNA extraction from the original tissue to eliminate the possibility of lab/handling errors. To assess the chance that matching pairs of samples were not the same individual, I implemented the recursive calculations of Chakraborty & Schull (1976) to compute the probability of identity, $P_{ID}(r,n)$, defined as the probability that two randomly selected individuals sharing pairwise relationship r would have only n or fewer loci with non-matching genotypes. Potential differences in iteroparity rate between hatchery programs and between sexes were explored using a z-test.

2.2.5 *Pedigree reconstruction*

Following release, juvenile steelhead from the Russian River migrate to the ocean and may then return to spawn at age two, three, or four. Therefore, I treated adult fish spawned in 2007, 2008 and 2009 as the potential parents of fish returning to spawn in 2009, 2010 and 2011 (Figure 2.2). I assigned parentage using the program SNPPIT (Anderson 2012), which identifies the most likely pair of parents for each offspring, then assesses the statistical confidence in the assignments using a novel, efficient simulation method, which is reported as a False Discovery Rate (FDR) score for that trio. For each offspring, the estimated FDR is what would be achieved if the parentage assignments of that offspring and of all offspring with more certain parentage were accepted. I assumed a genotyping error rate of 0.005 per gene copy, which corresponds to roughly 1% per locus, for most loci; however, based on Mendelian incompatibilities in reconstructed trios, I were able to estimate the genotyping error rate directly for 12 loci and I set rates accordingly (between 0.007 and 0.05 depending on the locus). I excluded fish with 10 or more missing loci (85 loci minimum) from the analysis.

I performed two runs of SNPPIT, with the first censoring all information about a fish's reported sex or day of spawning. Thus, for example, any pair of fish spawned in 2008 were potential mates, even if they were reported as the same sex or as spawned on different days. The second run included information about reported sex and spawn date of every fish. Comparison of these runs allowed some minor metadata errors to be rectified. In both runs, I selected a significance threshold so that the FDR was

near 0.005, such that I expect fewer than one of every 200 parentage assignments to be incorrect.

Finally, I assessed the accuracy of the reported FDR by performing an analysis in which the 2011WSH adult fish were treated as potential parents of the 2007WSH adults using the same parameters as in the other parentage runs. I expected to have zero parent/offspring trio assignments in this analysis.

2.2.6 *Age structure, reproductive success, spawning time*

Age of returning adults was determined for the 2007, 2008, and 2009 cohorts. Fish from the 2007 cohort could be identified when they returned at age two, three, and four in 2009, 2010 and 2011, respectively; fish from the 2008 cohort identified returning at age two and three, in 2010 and 2011, respectively; and fish from the 2009 cohort identified only at age two in 2011 (Figure 2.2). I compared the proportion of fish from the 2007 and 2008 cohorts returning at age two and three using a z-test, and assessed the age distribution of females and males throughout the spawn season.

The distribution of family sizes and number of mates per parent were calculated from the inferred parent-offspring trios for fish returning in years 2007, 2008, and 2009. This analysis included only those parents with at least one offspring found in the pedigree reconstruction. The number of mates per parent was not normally distributed, and could not be appropriately transformed, so a non-parametric Kruskal-Wallis test was used to evaluate differences both within years and between years (2007 and 2008) for female and male parents separately and Levene's test for homogeneity to compare variances.

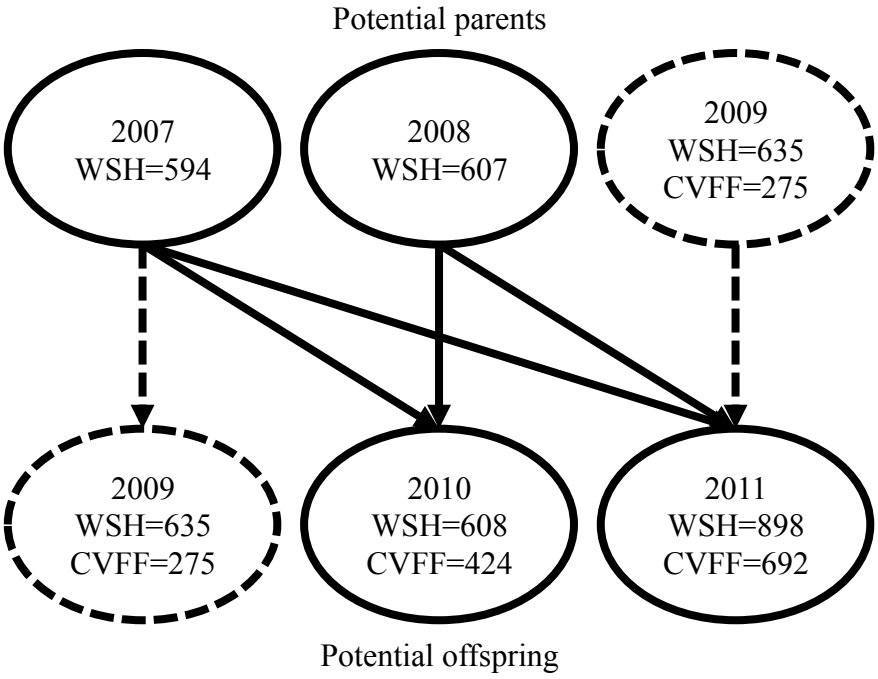

Figure 2.2: Graphical representation of the experimental design. Note that fish returning in 2009 (dashed circles) can be offspring of fish returning in 2007 as well as parents of fish returning in 2011. Numbers correspond to the total number of fish spawned that year at the two programs.

Reproductive success was estimated by counting the number of offspring per parent that returned to one of the two hatcheries. The number of offspring per parent was also not normally distributed, so a Kruskal-Wallis test was used to detect differences both between 2007 and 2008 and between female and male parents within years, separately. The variance in reproductive success between sexes was compared using Levene's test for homogeneity of variance across groups.

T-tests were used to compare a) the mean spawning day for age-two and age-three fish returning in 2010 and 2011, and for females and males separately, and b) the birth dates (parents' mean spawning day) for age-two and age-three fish. Age-four fish

were not included, because they were conclusively identified only in 2011. Since the pedigree analysis identified family relationships, heritability of the spawning day could be estimated. Returning fish enter the hatchery volitionally and are spawned on the date of their biological readiness, as determined by the hatchery staff. Heritability (h^2) was estimated as the slope of the parent-offspring regression line. Spawning days of all full-siblings within each family were averaged. The relationship between spawning day of the parents and the average spawning day of: a) all their offspring, b) their female offspring, and c) their male offspring were assessed separately, as were the relationships between the spawning day of the parents and the average spawning days of their age-two and age-three and older (3+) offspring. An analysis of covariance (ANCOVA) was used to test for differences in heritability between these groups. Heritability of spawning time was also evaluated using a linear regression between full siblings. Pairs of siblings from each family and pairs of presumably unrelated individuals were randomly selected 10,000 times with replacement. The slope of the regression line (*i.e.* heritability) for both groups was then compared using an ANCOVA. Heritability of spawning time was also estimated using a one-way analysis of variance on full-siblings.

All statistical analyses were performed using the R software (R Development Core Team 2011).

Table 2.2: Sample numbers for broodstock from the two programs. WSH, Warm Springs Hatchery; CVFF, Coyote Valley Fish Facility. Some individuals were spawned multiple times the same year. The total number of individuals indicates unique fish with ≥ 10 loci missing in their genotypes (see text).

Spawn year and program	Number of samples	Missing loci	Spawned twice	Spawned three times	Females	Males	Total no. of individuals
2007WSH	601	2	5	0	196	398	594
2008WSH	632	18	7	0	175	432	607
2009WSH	672	7	24	3	213	422	635
2010WSH	662	17	33	2	191	417	608
2011WSH	979	17	62	1	224	674	898
Total	3546	61	131	6	999	2343	3342
2009CVFF	283	0	8	0	107	168	275
2010CVFF	457	30	3	0	149	275	424
2011CVFF	710	10	8	0	210	482	692
Total	1450	40	19	0	466	925	1391

2.3 Results

Genotypes were collected from a total of 4,996 tissue samples (Table 2.2), including 3,546 from WSH and 1,450 from CVFF. Genotypes from 101 samples were excluded due to missing data ($\geq$ 10 missing loci), leaving 4,895 samples for further analyses. Some of these samples were duplicates from the same individual (see below).

2.3.1 *Matching samples and iteroparity rate*

The probability that two different, unrelated individuals would have identical genotypes at nearly all 96 loci was very small. In identifying matching genotypes, I allowed up to four mismatches to account for genotyping errors, but there were no pairs of genotypes that differed at more than two but less than five alleles. Given the allele frequencies in the populations, the probability of identity $P_{ID}(r,n)$, with $n = 4$, where n

is the number of loci with mismatches, was less than 10^{-23} for $r =$ "unrelated" and less than 10^{-10} for $r =$ "full sibling". Accordingly, samples with identical or nearly identical genotypes were considered to be from the same individual.

There were 150 individuals identified that were sampled and presumably spawned twice in the same year (Table 2.2). Of those, 54 were sampled more than once the same day and 96 were sampled on different days. In addition, six males were spawned three times in the same year. The proportion of fish sampled multiple times was strongly male biased (88.9%). I also identified 29 (0.89%) individuals that spawned in two different, always consecutive, years: two in 2007/2008, 12 in 2009/2010 and 15 in 2010/2011. Two iteroparous individuals returned to different facilities in the two years (one to WSH in 2009 and CVFF in 2010, and another to CVFF in 2010 and WSH in 2011). The total proportion of iteroparous individuals was 0.86% at WSH and 1.14% at CVFF ($z =$ -0.7, $p = 0.48$) and was similar for females and males over all years (females: 0.87%, males: 0.95%; $z =$ -0.24, $p = 0.84$).

2.3.2 *Sex determination*

Comparison of the phenotypic and genotypic sex determinations showed a proportion of 1.45% mismatches. These mismatches were resolved in two ways: 1) comparing the sex determinations for matching samples, and 2) by running the parentage analysis without including information about sex and identifying apparent same-sex pairs. Using the matching samples test I found six cases in which two samples identified as the same fish were assigned to different phenotypic but same genotypic sexes, and

one case in which the genotypic sexes were different but not the phenotypic. Parentage analysis without restricting mates to be of opposite sex identified four parent pairs with the same phenotypic but different genotypic sexes, suggesting that the phenotypic sex recorded was incorrect. On the other hand, one individual identified as phenotypic male and genotypic female was assigned to parent pairs with 4 different individuals, all identified phenotypically and genotypically as females. This suggests that the individual failed for the Y chromosome-linked gene and that the phenotypic identification is correct. In all cases, the phenotypic or genotypic sex assignments were corrected prior to the final parentage analysis.

2.3.3 *Pedigree reconstruction*

Two pedigree reconstruction analyses were performed: one in which mates were not constrained to be of opposite sex nor to have spawned on the same date, and the other in which they were. A total of 1,807 putative mother-father-offspring trios were identified in the analysis in which mates were not constrained to be of opposite sex nor to have spawned on the same date. Nineteen of the trios identified in this unconstrained analysis were not present in the constrained one. These 19 trios had a high FDR score (greater than 1%) and low maximum posterior probabilities. Two offspring were assigned a parent pair in both analyses but, in both cases, one parent was the same and the other was different. Both parent pairs were male-male in the unconstrained run and female-male in the constrained run. This ambiguity indicates that the assignment in the constrained run is likely to be correct and that a close relative

of the true mother is present in the dataset.

After reconciling the results from the constrained and unconstrained parentage analyses, the final number of offspring assigned a parent pair was 1,787. Among these, the mean posterior probability of the parent/offspring trio relationship was 0.9929 (range 0.497 - 0.999) and the mean FDR was 1.25×10^{-5} (range 0 - 0.00198). In this analysis, a FDR of 0.00198 indicates that no more than three to four parentage assignments out of the 1,787 are expected to be incorrect. This high confidence is supported by the analysis that treated the 2011WSH adult fish as potential parents of the 2007WSH adult fish, in which no false positive assignments were found.

The 1,787 reconstructed trios correspond to 50.59% of the adult offspring assigned to a parent pair, which was similar to expectations, since most of the fish from CVFF were born before sampling there began, as were some of the fish from WSH in 2009 (and likely also a few age-four fish in 2010) and because a small number of 2008WSH spawners were not sampled. This corresponds to 70.76% of the fish from WSH and 19.55% of the fish from CVFF with parental trios identified (Table 2.3). There were 15 fish born to CVFF parents that returned to spawn at WSH, while another 15 were from WSH parents but returned to CVFF, which yields estimates of migration rate of 0.99% for WSH fish and 5.51% for CVFF. This higher rate of migration from CVFF to WSH could be due to incubation at WSH of juvenile fish from CVFF parents, but also to the location of WSH lower in the basin.

I was able to identify four three-generation families from the pedigrees in which the offspring (one or two per family), the two parents, and all four grandparents are

Table 2.3: Parent pair assignments for progeny returning as adults in 2009, 2010 and 2011. Note that fish returning in 2009 can be progeny of fish spawned in 2007 and parents of fish returning in 2011. WSH, Warm Springs Hatchery; CVFF, Coyote Valley Fish Facility.

		Returning adults assigned to parents from:				
Spawn year and program	Total offspring	2007WSH	2008WSH	2009WSH	2009CVFF	Total assigned
2009WSH	635	231	-	-	-	231
2010WSH	608	376	152	-	-	528
2011WSH	898	11	277	453	15	756
Total	2141	618	430	453	15	1515
2009CVFF	275	1	-	-	-	1
2010CVFF	424	1	3	-	-	4
2011CVFF	692	0	10	0	257	267
Total	1391	2	13	0	257	272

known. I also found 100 three-generation families in which the offspring (mean = 1.72, range = 1 - 12 offspring per family), the two parents, and just one grandparent pair are known. Of these, there were 97 in which the paternal grandparents were identified, and three with just the maternal grandparents found, which is due to the younger mean age at reproduction for males (see below).

2.3.4 *Age structure of returning adults*

I assessed the age at first spawning for fish born in 2007, 2008, and 2009. From 620 fish assigned to parents spawned in 2007 (cohort 2007), a total of 232 (37.4%) returned at age two (8.2% females and 91.8% males), 377 (60.8%) returned at age three (43.2% females and 56.8% males), and only 11 (1.8%) at age four (63.6% females and 36.4% males). A total of 443 fish were assigned to 2008 parents (cohort 2008), with 155 (35.0%) age-two (3.2% females and 96.8% males) and 287 (64.8%) age-three offspring

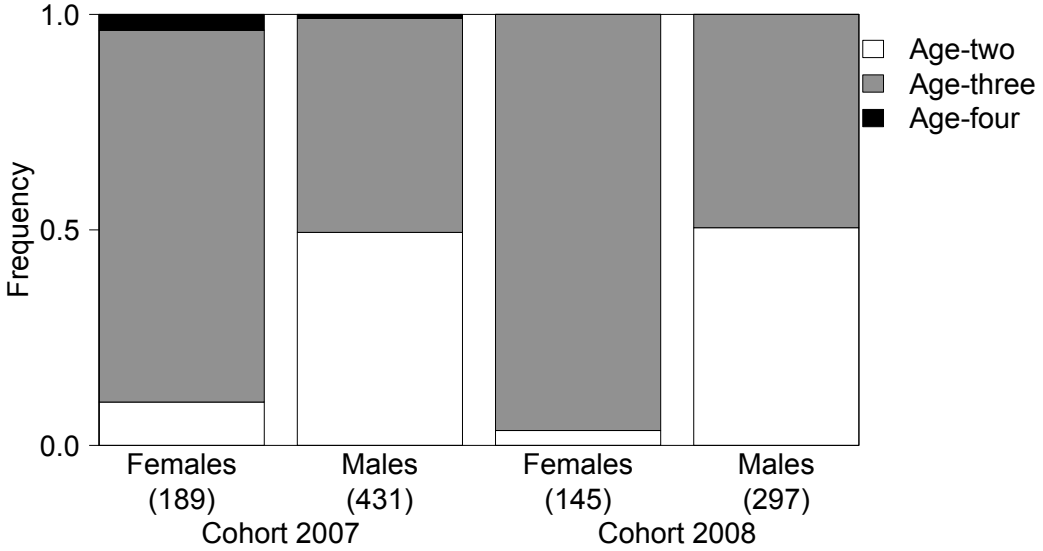

Figure 2.3: Age distribution of the returning adults by cohort (2007 and 2008) and by gender within each cohort. Numbers in parenthesis indicate the total number of fish. White bars represent age two, grey bars age three and black age four fish.

(48.8% females and 51.2% males; Figure 2.3). Note that age-four fish from cohort 2008 would not return during the study period. Finally, 725 age-two fish from cohort 2009 returned in 2011 (6.3% females and 93.7% males) to either WSH or CVFF. There was no difference in the proportion of fish that returned at age two ($z = 0.945$, $p = 0.344$) and age three ($z =$ -1.083, $p = 0.279$) in either cohort. The age distribution of all adults spawned from 2009 to 2011 revealed that 31.5% of the returning adults assigned to parents were spawned at age two, of which 6.3% were females and 93.7% were males.

Spawning time differed by age, with age-two fish spawning later in the season (mean spawn week = 9.55) than age-three fish (mean = 7.18) across all years ($t = 13.34$, $p < 0.001$). The same pattern was observed for females ($t = 15.55$, $p < 0.001$) and males ($t = 11.86$, $p < 0.001$) separately (Figure 2.4). In contrast, I observed that age-two fish

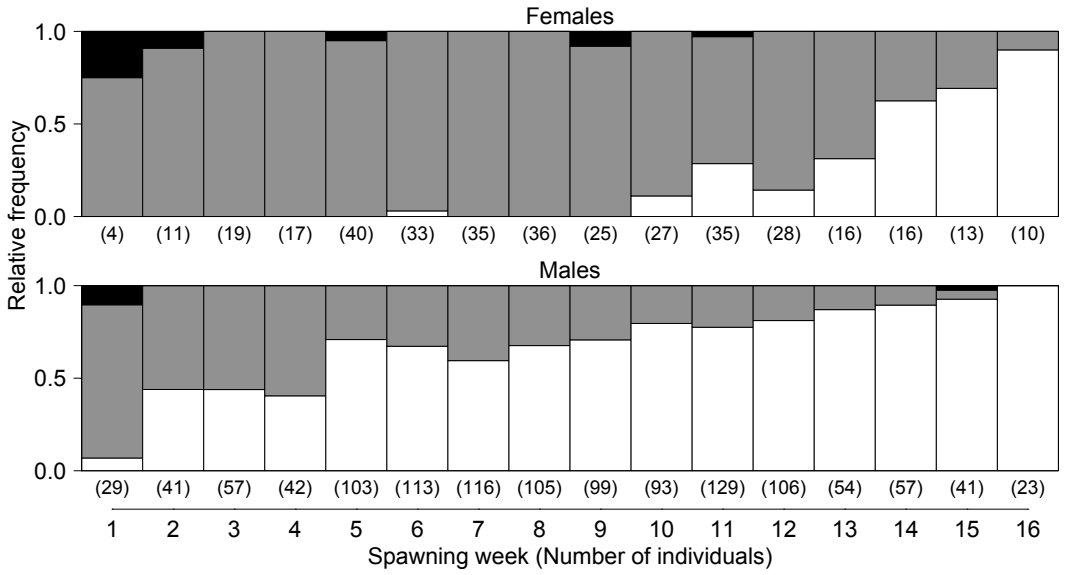

Figure 2.4: Age distribution throughout the spawn season of cohorts 2007 and 2008. Age four offspring are represented just in cohort 2007. Numbers in parenthesis indicate the total number of fish. White bars represent age two, grey bars age three and black age four fish.

tend to come from parents spawned earlier in the season than age-three fish, whose parents spawn later (t = -3.53, $p < 0.001$).

2.3.5 *Distribution of family sizes and reproductive success*

The 1,787 parent-offspring trios identified contained a total of 948 parent pairs, distributed in 295 pedigrees, and included 670 male parents and 504 female parents. The smallest pedigrees consisted of one offspring and its parents and accounted for 38.6% of all pedigrees, while the largest pedigree contained a total of 32 male parents, 20 female parents, and 76 offspring. The mean full-sibling family size amongst the returning adults was 2.0 offspring per parent pair (range 1-27; Figure 2.5).

Significant differences were found in the distribution and variance in number

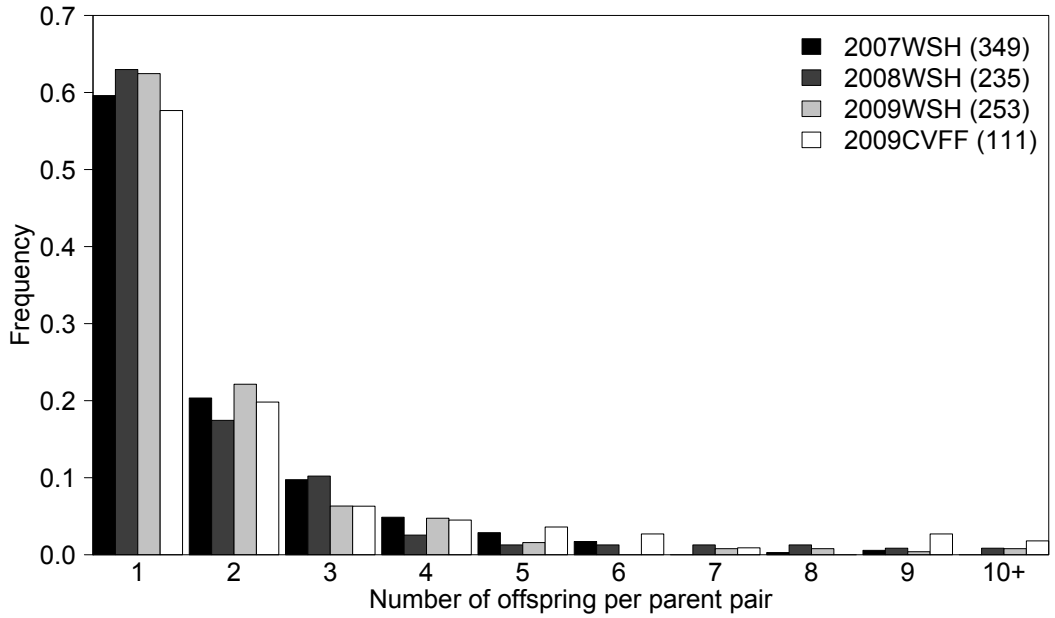

Figure 2.5: Relative frequency of offspring produced by a parent pair (full-sibling family size) per year and program. Numbers in parenthesis indicate number of parent pairs. Black bars represent pairs from 2007WSH, dark grey bars from 2008WSH, light grey bars from 2009WSH, and white from 2009CVFF.

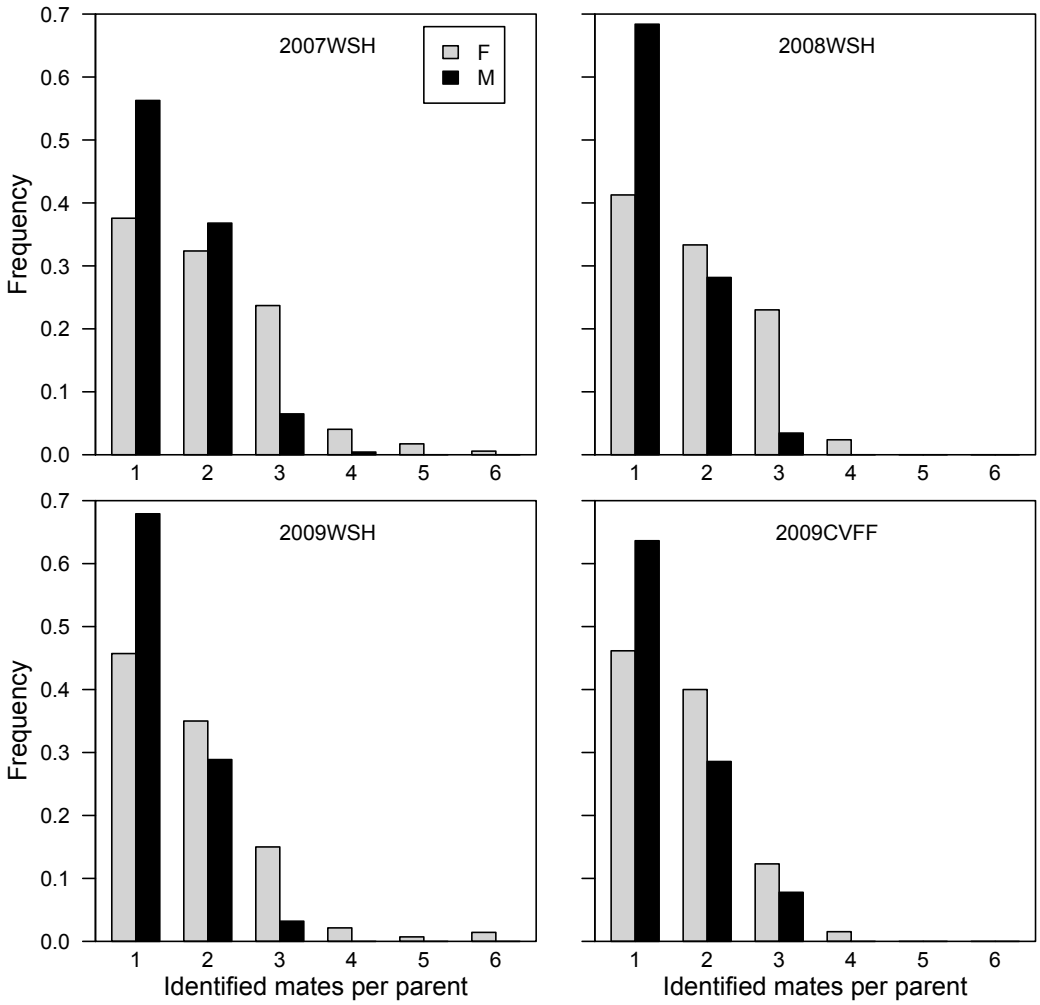

Figure 2.6: Relative frequency of identified mates per parent. Grey bars represent female parents and black bars correspond to male parents.

of mates per parent between females and males over all years, in which females had on average 1.88 mates (range 1-6) while males had 1.42 (range 1-4) ($\chi^2 = 81.42$, $p < 0.001$; and males$_{CV} = 0.42$, females$_{CV} = 0.5$, F = 71.91, $p < 0.001$). This trend was maintained when females and males were compared within years except for 2009CVFF parents when corrected for multiple tests (2007WSH: $\chi^2 = 27.24$, $p < 0.001$; 2008WSH: $\chi^2 = 30.68$, $p < 0.001$; 2009WSH: $\chi^2 = 22.34$, $p < 0.001$; 2009CVFF: $\chi^2 = 4.53$, $p = 0.033$; Figure 2.6). I found no differences in the number of mates per female parent between 2007 and 2008 ($\chi^2 = 1.02$, $p = 0.31$), but differences were observed between male parents ($\chi^2 = 6.67$, $p = 0.009$).

The mean number of returning offspring per male parent was 2.8 (range 1-51) and per female parent 3.6 (range 1-32). The male parent with the highest reproductive success (51 offspring) was from CVFF in 2009 and was spawned with two females, one of which had the highest reproductive success for a female (32 returning offspring) and produced the largest full-sibling family found (27 offspring). Females had higher mean reproductive success ($\chi^2 = 41.41$, $p < 0.001$), but a lower coefficient of variation (males$_{CV} = 1.17$, females$_{CV} = 0.99$; $F_{1,1172} = 8.54$, $p < 0.05$) than males over all three years (Figure 2.7), which is related to the fact that female's egg lots are always exposed to milt from more than one male in these programs. Significant differences were found between sexes within years except for 2009CVFF parents ($\chi^2 = 1.08$, $p = 0.298$). No significant differences in reproductive success between 2007 and 2008 were observed for either females or males.

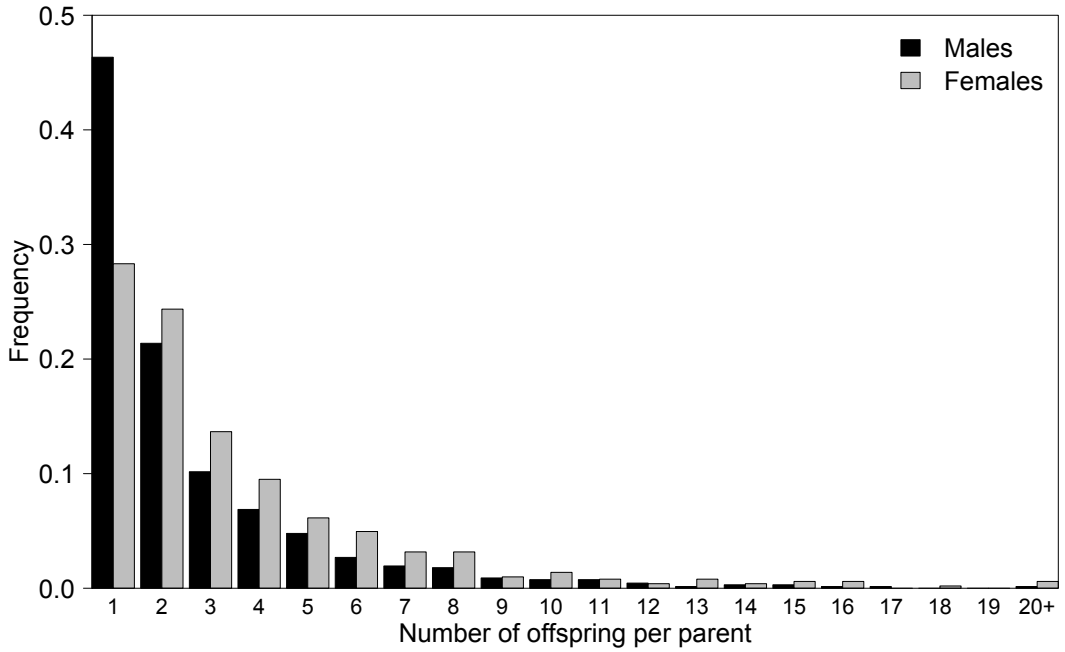

Figure 2.7: Relative frequency of offspring produced by a parent across all years. Black bars represent male parents and grey bars female parents.

2.3.6 *Heritability of spawning time*

I observed a strong positive correlation between the parent and offspring spawning day across all years ($F_{1,1089} = 510$, $R^2 = 0.319$, $p < 0.001$) and when female ($F_{1,281} = 132.6$, $R^2 = 0.32$, $p < 0.001$) and male ($F_{1,806} = 381.6$, $R^2 = 0.321$, $p < 0.001$) offspring were considered separately. Heritability (h^2) of spawning time was high for both female and male offspring combined ($h^2 = 0.512$), and for female ($h^2 = 0.563$) and male ($h^2 = 0.497$) offspring separately (Table 2.4 and Figure 2.8a), as well as when age-two ($h^2 = 0.545$) and age-3+ ($h^2 = 0.548$) offspring were considered separately. The ANCOVA indicated no statistically significant difference in heritability between any of these groups (Table 2.4).

Table 2.4: Heritability (h^2) of spawning time estimated by parent-offspring and full-sibling pair regression. Regression goodness of fit (R^2) and standard error (SE) are indicated. (*) Significance at $p < 0.05$ level.

	Parent-offspring					Full-sibs Random pairs
	All	Male	Female	Age 2	Age 3+	
h^2	0.512*	0.497*	0.563*	0.545*	0.548*	0.505*
R^2	0.319	0.321	0.32	0.387	0.434	0.253
SE	20.18	19.71	21.39	17.77	17.42	22.53

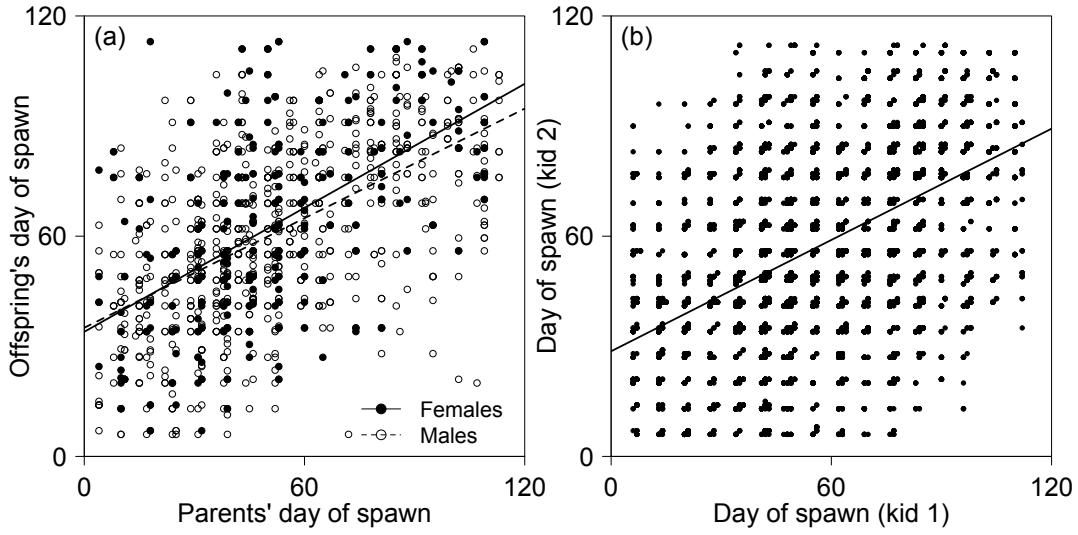

Figure 2.8: Heritability estimate of spawn date using two different methods. (a) Parent-offspring linear regression for females (closed circles, dashed line) and males (open circles, continuous line) offspring. (b) Full-siblings linear regression with 10,000 bootstraps.

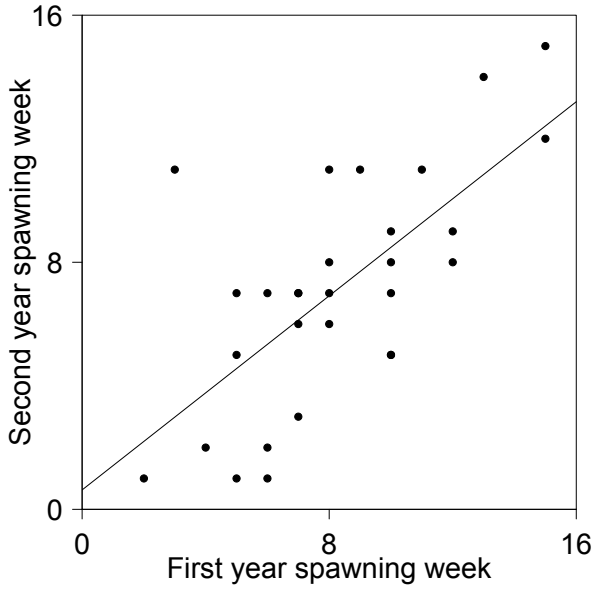

Figure 2.9: Correlation between the first and the second spawning week in iteroparus fish.

Heritability was also estimated from full-siblings using a similar linear regression analysis (Figure 2.8b) and pairs of randomly chosen individuals as a null distribution ($F_{1,9998} = 3381$, $R^2 = 0.253$, $p < 0.001$; Table 2.4). A strongly significant difference was found ($F = 1558.6$, $p < 0.001$) in the comparison between the regression line slopes (ANCOVA) of the random pairs and the full siblings. The ANOVA also found greater variation between than within families ($F_{1,2406} = 3016.2$, $p < 0.001$).

The spawning days of iteroparous fish in their first and second years of observed reproduction were strongly correlated ($R^2 = 0.447$, $p < 0.01$), further indicating the heritability of spawning time (Figure 2.9).

2.4 Discussion

Here, I use genetic pedigree analysis of steelhead, a fish that undertakes an ocean migration, to elucidate reproductive patterns and the basis for life history variation in one of the world's most widespread fish species. By sampling almost all reproducing adults over several years in two medium-sized hatchery programs in California, USA, I were able to employ large-scale parentage analysis with a modest number of SNP markers to identify the parents of most returning adults and infer several important and previously unknown aspects of the life history and biology of these important fish.

Pedigree reconstruction and parentage analysis using SNP markers have previously been conducted in humans (Delahunty *et al.* 1996) and livestock (Heaton *et al.* 2002; Rohrer *et al.* 2007), but only recently has their promise for the study of natural populations become evident (Pemberton 2008; Hauser *et al.* 2011). While statistical methodology for pedigree reconstruction and inference of relationships with genetic data has also long been in use (*e.g.* Marshall *et al.* 1998), the development of statistical algorithms (Anderson & Garza 2006) and software (Anderson 2012) that can handle large datasets and analyses effectively and efficiently have only recently become available. In conjunction with the increasing ease of development and genotyping of SNP assays for non-model organisms (Seeb *et al.* 2011), these advances portend a transition to pedigree-based methods employing SNP markers for many applications in ecological, evolutionary and conservation research.

The use of such parentage-based analyses as a surrogate for traditional tagging

methods also provided us with a large number of pedigrees, allowed us to evaluate patterns on a family level. This is exactly what is necessary for a classical estimation of trait heritability using parent/offspring and sibling/sibling regression (Fisher 1918). The finding that spawning time in steelhead is highly heritable is both a novel and important outcome of this approach, with implications for management and conservation of the species. The estimates of other life history parameters, including age distribution and family structure in the reproducing adults, provide an important baseline with which to evaluate the effects of hatchery production on natural populations of a species of conservation concern. Artificial propagation and subsequent supplementation can have numerous negative effects on natural populations (Utter 1998; Bryant & Reed 1999; Frankham 2008; Williams & Hoffman 2009; Christie *et al.* 2012) and detailed estimates of reproductive and behavioral trait values of the propagated population is a critical first step in understanding and minimizing these consequences.

2.4.1 *Sex determination*

While the sex determination of a fish in the field at the time of spawning seems trivial, there are many sources of error in recording and managing of the data, especially when handling data sets of thousands of individuals. I compared both phenotypic and genotypic sex determination to identify many of these errors and accurately assign the right sex to the fish. The results show the high accuracy of the SNP marker to determine sex. I found that the error rate is low (1.45%) based on the comparison with the visual identification. Also, I found that the combination of both methods for sex determination

is necessary to eliminate ambiguities and to assign fish accurately.

2.4.2 *Heritability of spawning time*

Timing of reproduction is crucially important for most organisms because, in a seasonally varying environment, it influences the conditions that the progeny will encounter (Brannon 1987; Reed *et al.* 2010). I demonstrate a strong genetic component to the time of spawning by steelhead. This high heritability was found for both males and females and using several methods. Numerous studies have examined the genetic architecture of life history traits of salmonids (*e.g.* Thrower *et al.* 2004; Leder *et al.* 2006; Nichols *et al.* 2008; Hecht *et al.* 2012), including spawning time (Siitonen & Gall 1989; Su *et al.* 1997; Quinn *et al.* 2000; Bentzen *et al.* 2001; Dickerson *et al.* 2005), but this is the first such examination in steelhead and the first using the classical pedigree-based approach in a population that is free-living for at least part of its life. Traits related to timing of reproduction (*i.e.* date of entry, maturation, and spawning) have shown the highest heritability values in *Oncorhynchus* species (Carlson & Seamons 2008). This suggests that there is strong selection pressure on these traits, which influences the performance and success of breeders. In addition, salmonids have strong natal homing behavior, which increases population genetic structure and local adaptation, such that differences in environmental conditions can affect life history traits in a modest number of generations. For example, recently introduced Chinook salmon populations in two New Zealand streams with very different environmental conditions rapidly evolved differences in the timing of migration, maturation, and breeding (Quinn

et al. 2000). Artificial selection due to environmental conditions in the hatchery can also play an important role in divergence of reproductive timing between hatchery fish and their natural progenitor population (Millenbach 1973; Quinn *et al.* 2002). This indicates that migration and reproductive timing may respond rapidly to selection and provide some ability to adapt in the face of changing environmental conditions.

Since age-two fish spawn later than older fish on average (Figure 2.4), and spawning time was found to be highly heritable in both ages (Table 2.4), I would expect age-two fish to be born later in the season, which is the opposite of the observed pattern; fish that spawn at age-two tend to come from early spawning parents and return to spawn as adults late in the season. One explanation could be that they require more time to mature, while age-three spawners are able to mature earlier in the season. This suggests either that age at maturity is not highly heritable in this population, that it is constrained by developmental requirements, or that it is overwhelmed by environmental conditions related to hatchery rearing.

2.4.3 *Iteroparity rate*

My observations of Russian River hatchery steelhead iteroparity are consistent with those previously reported by Hallock (1989) and Keefer *et al.* (2008) for the Sacramento and Columbia rivers, respectively, where just a few fish (less than 1%) returned a second time and none returned a third time. These studies also reported much lower iteroparity rates in hatchery- than in natural-origin fish, but I evaluated only hatchery-origin fish here. However, iteroparity in a naturally spawning steelhead

population in a proximate basin (the upper Eel River, ∼20km from CVFF), found a much higher rate (∼5%) of iteroparity (unpublished data), suggesting that hatchery production generally reduces repeat spawning in steelhead.

I found no difference between female and male iteroparity rates. This contrasts with reports for other salmonid species (Shearer 1992; Fleming 1998) and for natural steelhead populations (Ward & Slaney 1988; Keefer *et al.* 2008; Seamons & Quinn 2010) where female repeat spawners are more common than males. Male-male competition for mates in anadromous salmonids is more intense than that of female-female competition for nesting areas, which could reduce the post-spawning survival of males compared to females (Fleming & Gross 1994; Fleming 1996). Artificial spawning eliminates competition among males for mating opportunities, and could therefore increase male post-spawning survival and iteroparity rates.

2.4.4 *Pedigree reconstruction*

I reconstructed parent/offspring trios with high confidence, as reflected by high maximum posterior probabilities and low FDR scores. The proportion of fish with parent pairs identified was high at WSH (70.76% adult offspring assigned to parents) but low at CVFF (19.55%, Table 2.3). The unidentified parents can be explained by several factors, including the lack of sampling of parents from years prior to the study period, the removal from the analysis of some samples due to missing data, the lack of sampling on one spawn day in 2008, as well as the likely, but unreported, incorporation of some natural-origin fish into the broodstock in these programs. Migration of fish

from CVFF to WSH and their subsequent use as broodstock could also explain some of the unidentified parents in all years except 2011.

However, some fish are likely not assigned parents because of lack of statistical confidence, even though their parents' genotypes are available. I refer to this as a "false negative error." There is a direct negative relationship between the FDR and the false negative rate (Anderson & Garza 2006). Unfortunately, it is difficult to estimate the false negative rate from the genetic data alone, but I were able to use the 2011WSH spawners (most of whose parents should be represented in the genotyped samples) and associated information to estimate the false negative rate for this program (for details of this estimation see E. C. Anderson's supporting information in Abadía-Cardoso *et al.* 2013). My estimate of the false negative rate for this study was 10.3%. That is, if I have accounted for all sources of missing parental genotypes, then about 10% of the juveniles were not assigned a parent pair even though their parents were amongst the genotyped samples. However, if I have failed to account for only about 3% of the fish used as broodstock, which could result from incomplete sampling at the hatchery, loss of samples between spawning and dataset completion, an unusually high rate of migration from distant hatchery programs, or misidentified natural-origin fish, then the false negative rate would be close to zero.

2.4.5 *Age structure of returning adults*

The proportion of fish returning at age two and age three was similar for the 2007 and 2008 cohorts. However, significantly more males than females returned at age

two, whereas females more commonly returned at age three, which is consistent with results of a previous study of hatchery steelhead (Tipping 1991). It has been suggested that age at maturity is determined genetically for one sex and environmentally for the other (Ward & Slaney 1988; Tipping 1991). However, it is unclear to what extent this is true, and it is probable that a combination of genetic and environmental factors affect age of maturity in both sexes.

I observed a high proportion (about 30% on average) of age-two spawners at both WSH and CVFF. The proportion of age-two male spawners is even higher, exceeding 50% of male parents in 2009. This is in contrast to the management plan for these hatchery programs (FISHPRO 2004) that recommends less than 1% of spawners be age-two fish. While no age structure information is available for naturally spawning fish in the Russian River, the proportion of age-two spawning adults is much higher than what is generally seen in steelhead (Busby *et al.* 1996) and for natural-origin steelhead in proximate basins to both the north (Eel River; unpublished data) and the south (Waddell Creek; Shapovalov & Taft 1954), where the proportions of age-two returning adults were less than 5%. If age at maturity has a heritable component in this population, as has been shown in other salmonids (Carlson & Seamons 2008), then overrepresentation of age-two fish in the spawners, relative to the reproductive success that they would garner in natural spawning situations, will induce selection favoring earlier maturation. Substantial introgression by hatchery fish in the Russian River (Deiner *et al.* 2007) could then shift the age structure of naturally spawning populations. This would have consequences for reproductive success and fitness, especially in females,

since younger females are smaller than older females and size is strongly correlated with female fecundity in steelhead (Shapovalov & Taft 1954).

2.4.6 *Distribution of family sizes and reproductive success*

In both programs, management goals specify that female spawners are to be crossed with two or three males, whereas males are to be crossed with only one female, except when there are not a sufficient number of males to cross every female with unique males, which should be a relatively rare occurrence. The results indicate that reuse of males is common. This is in evidence both in the matching sample analysis as well as with the finding that 36.6% of males from the two programs that produced returning offspring did so with more than one female. Much of this reuse is on the same spawn day, but the identification of some fish spawned more than once in different spawning weeks indicates that they were spawned, released downstream of the hatchery and then reentered and were spawned again. The reuse of males will reduce effective population size relative to a crossing scheme where every male contributes only once, regardless if single-pair, promiscuous, or factorial mating is used. A promiscuous crossing scheme, in which multiple males are mated with each female and most are used only once, as in these programs, is expected to increase genetic diversity and the number of families and reduce the chance of inbreeding relative to single-pair or factorial mating (Pearse & Anderson 2009). A promiscuous breeding strategy, in which both females and males breed with multiple partners, has been observed in natural populations of steelhead (Shapovalov & Taft 1954; Seamons *et al.* 2004). The use of a genetic pedigree-based

monitoring method will allow routine evaluation of the effective population size and the distribution of mating partners in species with multiple bouts of reproduction.

Similar reproductive success for males between years and for females between years was observed, but females had a higher average number of offspring and smaller coefficient of variation than males over all years. This is not surprising, because females are always crossed with more than one male in these programs and, assuming that females and males have the same probability of surviving to complete an ocean migration and return to spawn, females were expected to have more offspring on average than males. However, if no offspring were allocated to a particular parent, it does not necessarily mean that the parent did not produce returning offspring, since not all returning adults are used as broodstock. In addition, offspring do not always return to the hatchery and instead spawn in natural areas, and some offspring genotypes were discarded due to missing data.

I demonstrate here that large-scale parentage inference with SNP markers is an effective tagging method for a species that spends most of its life in the ocean before returning to reproduce in freshwater. This innovative intergenerational genetic tagging method holds great promise for the study of high-fecundity organisms, because juveniles are not handled until and unless they survive the high-mortality portions of their life history. The associated pedigrees are an additional valuable resource with many potential uses. For example, in species subject to artificial propagation, they can be used to understand the effects of breeding programs on quantitative genetic traits. Such information will allow formulation of better strategies for supplementation

programs and ultimately lead to more effective conservation and management plans.

2.5 References

Abadía-Cardoso A, Anderson EC, Pearse DE, Garza JC (2013) Large-scale parentage analysis reveals reproductive patterns and heritability of spawn timing in a hatchery population of steelhead (*Oncorhynchus mykiss*). Molecular Ecology 22, 4733-4746.

Abadía-Cardoso A, Clemento AJ, Garza JC (2011) Discovery and characterization of single nucleotide polymorphisms in steelhead/rainbow trout, *Oncorhynchus mykiss*. Molecular Ecology Resources 11, 31-49.

Aguilar A, Garza JC (2008) Isolation of 15 single nucleotide polymorphisms from coastal steelhead, *Oncorhynchus mykiss* (Salmonidae). Molecular Ecology Resources 8, 659-662.

Anderson EC (2012) Large-scale parentage inference with SNPs: an efficient algorithm for statistical confidence of parent pair allocations. Statistical Applications in Genetics and Molecular Biology 11, 12.

Anderson EC, Garza JC (2006) The power of single-nucleotide polymorphisms for large-scale parentage inference. Genetics 172, 2567-2582.

Araki H, Berejikian BA, Ford MJ, Blouin MS (2008) Fitness of hatchery-reared salmonids in the wild. Evolutionary Applications 1, 342-355.

Araki H, Cooper B, Blouin MS (2007) Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. Science 318, 100-103.

Augerot X (2005) Atlas of Pacific salmon: The first map-based status assessment of salmon in the North Pacific. University of California Press, Berkeley, CA.

Avise JC, Jones AG, Walker D, DeWoody JA (2002) Genetic mating systems and reproductive natural histories of fishes: lessons for ecology and evolution. Annual Review of Genetics 36, 19-45.

Baskett ML, Waples RS (2013) Evaluating alternative strategies for minimizing unintended fitness consequences of cultured individuals on wild populations. Conservation Biology 27, 83-94.

Bentzen P, Olsen JB, McLean JE, Seamons TR, Quinn TP (2001) Kinship analysis of Pacific salmon: Insights into mating, homing, and timing of reproduction. Journal of Heredity 92, 127-136.

Blouin MS (2003) DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. Trends in Ecology & Evolution 18, 503-511.

Brannon EL (1987) Mechanisms stabilizing salmonid fry emergence timing. Canadian Special Publication of Fisheries and Aquatic Sciences 96, 120-124.

Brunelli JP, Wertzler KJ, Sundin K, Thorgaard GH (2008) Y-specific sequences and polymorphisms in rainbow trout and Chinook salmon. Genome 51, 739-748.

Bryant EH, Reed DH (1999) Fitness decline under relaxed selection in captive populations. Conservation Biology 13, 665-669.

Busby PJ, Wainwright TC, Bryant GJ, *et al.* (1996) Status review of west coast steelhead from Washington, Idaho, Oregon, and California. NOAA Technical Memorandum NMFS-NWFSC-27. US Department of Commerce.

Campbell NR, Overturf K, Narum SR (2009) Characterization of 22 novel single nucleotide polymorphism markers in steelhead and rainbow trout. Molecular Ecology Resources 9, 318-322.

Carlson SM, Seamons TR (2008) A review of quantitative genetic components of fitness in salmonids: implications for adaptation to future change. Evo. App. 1, 222-238.

Chakraborty R, Schull W (1976) A note on the distribution of the number of exclusions to be expected in paternity testing. American Journal of Human Genetics 28, 615-618.

Champagnon J, Elmberg J, Guillemain M, Gauthier-Clerc M, Lebreton JD (2012) Conspecifics can be aliens too: a review of effects of restocking practices in vertebrates. Journal for Nature Conservation 20, 231-241.

Christie MR, Marine ML, French RA, Blouin MS (2012) Genetic adaptation to captivity can occur in a single generation. Proceedings of the National Academy of Sciences of the United States of America 109, 238-242.

Clemento AJ, Abadía-Cardoso A, Starks HA, Garza JC (2011) Discovery and characterization of single nucleotide polymorphisms in Chinook salmon, *Oncorhynchus tshawytscha*. Molecular Ecology Resources 11, 50-66.

Colihueque N, Cardenas R, Ramirez L, Estay F, Araneda C (2010) Analysis of the association between spawning time QTL markers and the biannual spawning behavior in rainbow trout (*Oncorhynchus mykiss*). Genetics and Molecular Biology 33, 578-582.

Cooke SJ, Hinch SG, Wikelski M, *et al.* (2004) Biotelemetry: a mechanistic approach to ecology. Trends in Ecology & Evolution 19, 334-343.

Cooper JC, Scholz AT (1976) Homing of artificially imprinted steelhead (rainbow) trout, *Salmo gairdneri*. Journal of the Fisheries Research Board of Canada 33, 826-829.

Deiner K, Garza JC, Coey R, Girman DJ (2007) Population structure and genetic diversity of trout (*Oncorhynchus mykiss*) above and below natural and man-made barriers in the Russian River, California. Conservation Genetics 8, 437-454.

Delahunty C, Ankener W, Deng Q, Eng J, Nickerson DA (1996) Testing the feasibility of DNA typing for human identification by PCR and an oligonucleotide ligation assay. American Journal of Human Genetics 58, 1239-1246.

Dickerson BR, Willson MF, Bentzen P, Quinn TP (2005) Heritability of life history and morphological traits in a wild pink salmon population assessed by DNA parentage analysis. Transactions of the American Fisheries Society 134, 1323-1328.

Fisher RA (1918) The correlation between relatives on the supposition of Mendelian inheritance. Philosophical Transactions of the Royal Society of Edinburgh 52, 399-433.

FISHPRO (2004) Hatchery and genetic management plans for Russian River fish production facilities Coho salmon and steelhead., p. 73. FISHPRO, a Division of HDR and ENTRIX, inc. Available from: nrm.dfg.ca.gov.

Fleming IA (1996) Reproductive strategies of Atlantic salmon: ecology and evolution. Reviews in Fish Biology and Fisheries 6, 379-416.

Fleming IA (1998) Pattern and variability in the breeding system of Atlantic salmon (*Salmo salar*), with comparisons to other salmonids. Canadian Journal of Fisheries and Aquatic Sciences 55, 59-76.

Fleming IA, Gross MR (1994) Breeding competition in a Pacific salmon (Coho: *Oncorhynchus kisutch*): measures of natural and sexual selection. Evolution 48, 637-657.

Frankham R (2008) Genetic adaptation to captivity in species conservation programs. Molecular Ecology 17, 325-333.

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to conservation genetics. Cambridge University Press, Cambridge, UK.

Garrigue C, Dodemont W, Steel D, Baker CS (2004) Organismal and 'gametic' capture-recapture using microsatellite genotyping confirm low abundance and reproductive autonomy of humpback whales on the wintering grounds of New Caledonia. Marine Ecology-Progress Series 274, 251-262.

Glaubitz JC, Rhodes OE, Dewoody JA (2003) Prospects for inferring pairwise relationships with single nucleotide polymorphisms. Molecular Ecology 12, 1039-1047.

Haidle L, Janssen JE, Gharbi K, *et al.* (2008) Determination of quantitative trait loci (QTL) for early maturation in rainbow trout (*Oncorhynchus mykiss*). Marine Biotechnology 10, 579-592.

Hallock RJ (1989) Upper Sacramento River steelhead, (*Oncorhynchus mykiss*), 1952-1988. A report to the Fish and Wildlife Service., p. 85. Sacramento, California.

Hankin DG, Clark JH, Deriso RB, *et al.* (2005) Report of the expert panel on the future of the coded wire tag program for Pacific salmon: PSC Technical Report, p. 243. Available from: http://www.psc.org/info_codedwiretagreview.htm.

Hansen MM, Olivieri I, Waller DM, Nielsen EE (2012) Monitoring adaptive genetic responses to environmental change. Molecular Ecology 21, 1311-1329.

Hauser L, Baird M, Hilborn R, Seeb LW, Seeb JE (2011) An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye salmon (*Oncorhynchus nerka*) population. Molecular Ecology Resources 11, 150-161.

Heaton MP, Harhay GP, Bennett GL, *et al.* (2002) Selection and use of SNP markers for animal identification and paternity analysis in US beef cattle. Mammalian Genome 13, 272-281.

Hecht BC, Thrower FP, Hale MC, Miller MR, Nichols KM (2012) Genetic architecture of migration-related traits in rainbow and steelhead trout, *Oncorhynchus mykiss*. G3-Genes Genomes Genetics 2, 1113-1127.

Hendry AP, Day T (2005) Population structure attributable to reproductive time: isolation by time and adaptation by time. Molecular Ecology 14, 901-916.

Hoffnagle TL, Carmichael RW, Frenyea KA, Keniry PJ (2008) Run timing, spawn timing, and spawning distribution of hatchery- and natural-origin spring Chinook salmon in the Imnaha River, Oregon. North American Journal of Fisheries Management 28, 148-164.

Hudy M, Coombs JA, Nislow KH, Letcher BH (2010) Dispersal and within-stream spatial population structure of brook trout revealed by pedigree reconstruction analysis. Transactions of the American Fisheries Society 139, 1276-1287.

Keefer ML, Wertheimer RH, Evans AF, Boggs CT, Peery CA (2008) Iteroparity in Columbia River summer-run steelhead (*Oncorhynchus mykiss*): implications for conser-

vation. Canadian Journal of Fisheries and Aquatic Sciences 65, 2592-2605.

Leder EH, Danzmann RG, Ferguson MM (2006) The candidate gene, Clock, localizes to a strong spawning time quantitative trait locus region in rainbow trout. Journal of Heredity 97, 74-80.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7, 639-655.

Martínez A, Garza JC, Pearse DE (2011) A microsatellite genome screen identifies chromosomal regions under differential selection in steelhead and rainbow trout. . Transactions of the American Fisheries Society 140, 829-842.

Millenbach C (1973) Genetic selection of steelhead trout for management purposes. International Atlantic Salmon Symposium 4, 253-257.

Miller MR, Brunelli JP, Wheeler PA, *et al.* (2012) A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. Molecular Ecology 21, 237-249.

Nichols KM, Edo AF, Wheeler PA, Thorgaard GH (2008) The genetic basis of smoltification-related traits in *Oncorhynchus mykiss*. Genetics 179, 1559-1575.

NOAA (2006) Endangered and threatened species: final listing determinations for 10 distinct population segments of west coast steelhead (ed. NOAA), pp. 833-862. US Federal Register.

Palkovacs EP, Kinnison MT, Correa C, Dalton CM, Hendry AP (2012) Fates beyond traits: ecological consequences of human-induced trait change. Evolutionary Applications 5, 183-191.

Palsbøll PJ (1999) Genetic tagging: contemporary molecular ecology. Biological Journal of the Linnean Society 68, 3-22.

Pearse DE, Anderson EC (2009) Multiple paternity increases effective population size. Molecular Ecology 18, 3124-3127.

Pearse DE, Eckerman CM, Janzen FJ, Avise JC (2001) A genetic analogue of 'mark-recapture' methods for estimating population size: an approach based on molecular parentage assessments. Molecular Ecology 10, 2711-2718.

Pemberton JM (2008) Wild pedigrees: the way forward. Proceedings of the Royal So-

ciety B-Biological Sciences 275, 613-621.

Planes S, Jones GP, Thorrold SR (2009) Larval dispersal connects fish populations in a network of marine protected areas. Proceedings of the National Academy of Sciences of the United States of America 106, 5693-5697.

Quinn TP (2004) The behavior and ecology of Pacific salmon and trout. University of Washington Press, Seattle, WA.

Quinn TP, Peterson JA, Gallucci VF, Hershberger WK, Brannon EL (2002) Artificial selection and environmental change: countervailing factors affecting the timing of spawning by Coho and Chinook salmon. Trans of the Am Fish Soc 131, 591-598.

Quinn TP, Unwin MJ, Kinnison MT (2000) Evolution of temporal isolation in the wild: genetic divergence in timing of migration and breeding by introduced Chinook salmon populations. Evolution 54, 1372-1385.

R Development Core Team (2011) A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Reed TE, Waples RS, Schindler DE, Hard JJ, Kinnison MT (2010) Phenotypic plasticity and population viability: the importance of environmental predictability. Proceedings of the Royal Society B-Biological Sciences 277, 3391-3400.

Rohrer GA, Freking BA, Nonneman D (2007) Single nucleotide polymorphisms for pig identification and parentage exclusion. Animal Genetics 38, 253-258.

Seamons TR, Bentzen P, Quinn TP (2004) The mating system of steelhead, *Oncorhynchus mykiss*, inferred by molecular analysis of parents and progeny. Environmental Biology of Fishes 69, 333-344.

Seamons TR, Quinn TP (2010) Sex-specific patterns of lifetime reproductive success in single and repeat breeding steelhead trout (*Oncorhynchus mykiss*). Behavioral Ecology and Sociobiology 64, 505-513.

Seeb JE, Carvalho G, Hauser L, *et al.* (2011) Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in non-model organisms. Molecular Ecology Resources 11, 1-8.

Shapovalov L, Taft AC (1954) The life histories of the steelhead rainbow trout (*Salmo gairdneri gairdneri*) and silver salmon (*Oncorhynchus kisutch*) with special reference to Waddell Creek, California, and recommendations regarding their management. In: California Department of Fish and Game Fish Bulletin, 303 p.

Shearer WM (1992) The Atlantic salmon: natural history, exploitation and future management fishing. News Books, Oxford, U.K.

Siitonen L, Gall GAE (1989) Response to selection for early spawn date in rainbow trout, *Salmo gairdneri*. Aquaculture 78, 153-161.

Smoker WW, Gharrett AJ, Stekoll S (1998) Genetic variation of return date in a population of pink salmon: a consequence of fluctuating environment and dispersive selection? Alaska Fishery Research Bulletin 5, 46-54.

Su GS, Liljedahl LE, Gall GAE (1997) Genetic and environmental variation of female reproductive traits in rainbow trout (*Oncorhynchus mykiss*). Aquaculture 154, 115-124.

Swinnerton KJ, Groombridge JJ, Jones CG, Burn RW, Mungroo Y (2004) Inbreeding depression and founder diversity among captive and free-living populations of the endangered pink pigeon *Columba mayeri*. Animal Conservation 7, 353-364.

Thrower FP, Hard JJ, Joyce JE (2004) Genetic architecture of growth and early life-history transitions in anadromous and derived freshwater populations of steelhead. Journal of Fish Biology 65, 286-307.

Tipping JM (1991) Management briefs: heritability of age at maturity in steelhead. North American Journal of Fisheries Management 11, 105-108.

Utter F (1998) Genetic problems of hatchery-reared progeny released into the wild, and how to deal with them. Bulletin of Marine Science 62, 623-640.

Ward BR, Slaney PA (1988) Life history and smolt-to-adult survival of Keogh River steelhead trout (*Salmo gairdneri*) and the relationship to smolt size. Canadian Journal of Fisheries and Aquatic Sciences 45, 1110-1122.

Weitkamp LA, Wainwright TC, Bryant GJ, *et al.* (1995) Status review of Coho salmon from Washington, Oregon, and California. NOAA Technical Memorandum, NMFS-NWFSC-24.

Williams SE, Hoffman EA (2009) Minimizing genetic adaptation in captive breeding programs: a review. Biological Conservation 142, 2388-2400.

Wilson AJ, McDonald G, Moghadam HK, Herbinger CM, Ferguson MM (2003) Marker-assisted estimation of quantitative genetic parameters in rainbow trout, *Oncorhynchus mykiss*. Genetical Research 81, 145-156.

Chapter 3

Phylogeographic history of Pacific trout at the extreme southern end of their native range and risk of introgression from exotic hatchery trout

Abstract

Salmonid fishes are cold water piscivores with a native distribution spanning nearly the entire northern hemisphere. Trout in the genus *Oncorhynchus* are the most widespread of the salmonid fishes and also among the most important fish species in the world, due to their extensive use in aquaculture and extremely valuable fisheries. The trout that inhabit northwestern Mexico are the southernmost native salmonid populations in the world, and the least known in North America. They are unfortunately also facing serious threats to their continued existence. Previous work has described one new species, the Mexican golden trout (*O. chrysogaster*), and one new subspecies, Nelson's trout (*O. mykiss nelsoni*) in Mexico, but preliminary genetic analyses indicate that there is vastly more biodiversity in this group than formally described. Here I conducted a comprehensive genetic analysis of this important group of fishes by using novel genetic markers and techniques to elucidate the biodiversity of trout inhabiting northwestern Mexico, compared it to that of other species of Pacific trout, evaluate hypotheses regarding their evolutionary history, and measure introgression from non-native hatchery rainbow trout. This study revealed significant divergence between Mexican trout and the other species. I confirmed the vast genetic diversity present in the Mexican trout complex and the extremely strong genetic differentiation, not only between basins, but also at a smaller scale. I also found that introgression from non-native rainbow trout is present, but the genetic integrity of native trout is still maintained in many watersheds. This information will help to guide effective conservation strategies for this globally

important group of fishes.

3.1 Introduction

The first step in construction of an effective conservation strategy for any taxon is to document the diversity of biological units in that taxon and gain understanding of the evolutionary processes that result in the generation of those biological units (Mayden & Wood 1995). The taxonomic status of native trout inhabiting northwestern Mexico has been the subject of speculation and controversy for decades. Behnke (2002) considered this group as "the most diverse and the least known trout of western North America". Only two taxa from the Mexican trout complex have been formally described: Nelson's trout, *Oncorhynchus mykiss nelsoni* (Evermann 1908), distributed in the Río Santo Domingo in northern Baja California; and the Mexican golden trout, *O. chrysogaster* (Needham & Gard 1964) from ríos Fuerte, Sinaloa and Culiacán in the central highlands of the Sierra Madre Occidental (SMO). Both taxa are currently protected by Mexican law (SEMARNAT 2000) and the Mexican golden trout has been listed as Vulnerable by the IUCN since 1990 (IUCN 2010). Other formally undescribed groups of trout are found throughout the Sierra Madre Occidental and trout have been documented in four basins north of the range of the Mexican golden trout (NSMO here after): ríos Mayo, Yaqui, Guzmán, and Conchos, and in five other basins to the south (SSMO here after): ríos San Lorenzo, Piaxtla, Presidio, Baluarte, Acaponeta (Behnke 2002; Hendrickson *et al.* 2002, 2006; Ruiz-Campos *et al.* 2003), and potentially as far

south as Río San Pedro Mezquital (Espinoza-Pérez, pers. comm.). These groups have been considered as undescribed subspecies of *O. mykiss*, but, no conclusive evidence has been provided (Nielsen & Sage 2001; Behnke 2002).

The fossil record indicates that trout inhabited Mexico during the Pleistocene. The southernmost record for a fish assigned to the family Salmonidae is from the Lake Chapala, Jalisco, Mexico region near 20° North latitude (Cavender & Miller 1982). It has been suggested that the Gulf of California acted as a refugium for anadromous *O. mykiss* during the Pleistocene glaciations (Behnke 1992). These trout migrated from the Gulf into northwestern Mexico, Arizona and New Mexico. The subsequent increase in both ocean and river water temperatures constrained these trout to the high elevation headwaters of different river systems. Long isolation times gave rise to the Gila (*O. gilae*), Apache (*O. apache*), Mexican golden (*O. chrysogaster*) and, presumably, the other SMO trout.

For years, researchers have tried to investigate the evolutionary relationships among the Mexican trout complex using a variety of methods including karyology (Phillips & Rab 2001), morphology (Behnke 1992; Ruiz-Campos & Pister 1995; Ruiz-Campos *et al.* 2003), and genetic analyses with mitochondrial DNA (mtDNA; Nielsen *et al.* 1998; Camarena-Rosales *et al.* 2007) and microsatellites (Nielsen & Sage 2001; De los Santos-Camarillo 2008), but most of them focused on just one species (*i.e. O. chrysogaster*) or a few populations in specific regions (*i.e.* Río Yaqui or Río Mayo trout).

In a morphological variation analysis using Mahalanobis' distances that includes several populations of the Mexican trout complex, Ruiz-Campos *et al.* (2003)

found two main groups. The first group included *O. mykiss nelsoni* and the second group comprised trout inhabiting the SMO (*O. chrysogaster* and the other SMO trout). This last group was divided into four subgroups: *O. chrysogaster* was separated into two subgroups: 1) *O. chrysogaster* from Río Sinaloa, and 2) *O. chrysogaster* from ríos Fuerte and Culiacán along with trout from Río Piaxtla; 3) trout from ríos San Lorenzo, Baluarte, and Acaponeta; and 4) trout from ríos Yaqui and Mayo.

The first molecular study to include a population of Mexican trout in a phylogenetic analysis was that of Loudenslager *et al.* (1986), who used data from 36 allozyme loci to elucidate relationships between *O. gilae*, *O. apache*, *O. mykiss*, *O. clarkii* and trout from the Río Mayo only. They found that *O. gilae*, *O. apache* and Río Mayo trout showed a greater genetic identity to *O. mykiss* than to *O. clarkii*. They also reported that *O. gilae* and *O. apache* were sister taxa, and that Río Mayo trout were more closely related to *O. mykiss* than to any other species analyzed.

Nielsen *et al.* (1997) examined nominal *O. mykiss* from 15 California and two Mexican populations (*O. m. nelsoni* and Río Yaqui trout) with the control region of mtDNA and three nuclear microsatellite loci (Omy77, Omy207 and Ssa289). They concluded that *O. m. nelsoni* was closely related to Little Kern golden trout (*O. m. whitei*) and that Río Yaqui trout were considerably different than the rest of the populations analyzed. Additionally, Nielsen *et al.* (1998) analyzed the phylogenetic relationships between two species of Pacific salmon (*O. tshawytscha* and *O. kisutch*), four subspecies of cutthroat trout, *O. gilae*, *O. apache*, nine subspecies of *O. mykiss* (including *O. m. nelsoni*), and one population of Río Yaqui trout. The phylogenetic analysis showed well-

supported differentiation between species and was consistent with previous work, but the single mtDNA control region marker employed was not sufficiently informative for resolution at the subspecies level. Interestingly, Nielsen *et al.* (1998) also found a large deletion in the right-domain of the mtDNA control region in the Río Yaqui trout. This deletion had only been reported previously in humans and is related to mitochondrial disorders (Moraes *et al.* 1991), although no signs of such disorders have been reported in trout.

Camarena-Rosales (2007) evaluated restriction fragment length polymorphisms in one mtDNA region, including samples from most of the basins in northwestern Mexico where trout have been reported. The analysis divided the Mexican trout into four groups: 1) *O. m. nelsoni*; 2) *O. chrysogaster* subdivided into two groups; 3) trout from Río Piaxtla; and 4) ríos Mayo and Yaqui trout.

Microsatellite loci have seen widespread use in the study of *O. mykiss* population structure and interactions among different groups. There are two studies that have evaluated the genetic diversity in more than one population from northwestern Mexico using these markers. Nielsen and Sage (2001) evaluated 11 microsatellites in trout from ríos Yaqui, Mayo, and Guzmán, as well as *O. chrysogaster*. They showed a strong differentiation of Río Yaqui trout from the Mexican golden trout and population structure within the Yaqui basin.

The most comprehensive study to date using microsatellites included data from Nielsen and Sage (2001), as well as populations farther south (ríos San Lorenzo, Piaxtla, Presidio, Baluarte and Acaponeta). In this study, the presence of seven taxonomic

units inhabiting the SMO (in addition to *O. chrysogaster*) was proposed (De los Santos-Camarillo 2008).

The studies described above provide valuable insight into the high diversity of trout inhabiting northwestern Mexico, but it is evident that the incomplete sampling effort, small sample sizes, and the low resolution of these analyses have left many unresolved questions. There is a dire need for more information on the biodiversity and taxonomic status of trout in Mexico, as they are the southernmost populations of salmonid fish in their native range (the northern hemisphere), and are the only fish in this group that inhabit subtropical waters. Given the importance of trout in global aquaculture and fisheries, adaptation to such conditions is a critical trait to understand in these primarily cold-water fishes.

Conservation of the trout inhabiting northwestern Mexico first requires complete documentation of the genetic diversity of this species complex, as well as a complete understanding of the evolutionary history of these trout, which requires analysis of fish from all the Mexican basins in which native trout have been reported and a comparison with other trout species (*O. mykiss* and *O. clarkii*), specifically those with a presumably similar evolutionary history (*O. gilae* and *O. apache*).

The unique gene pool that is represented by these taxa is likely to go extinct due to threats by anthropogenic factors (*e.g.* habitat loss, logging, pollution and global climate change) without urgent documentation and conservation action. Moreover, the practice of introducing exotic hatchery rainbow trout (*O. mykiss irideus*) has caused them to be established in several drainages where native trout also occur (De los Santos-

Camarillo 2008). Several studies of California trout have reported introgression from genetically depauperate hatchery rainbow trout into wild populations, and this has become a substantial threat to native trout (Garza & Pearse 2008; Clemento *et al.* 2009). Hybridization of introduced rainbow trout with other native trout species has also been documented in the United States. One of the most-studied cases is that of the westslope cutthroat (*O. clarkii lewisi*) and exotic rainbow trout (Leary *et al.* 1985; Rubidge *et al.* 2001; Weigel *et al.* 2003; Allendorf *et al.* 2004). Several populations of westslope cutthroat trout have shown high degrees of introgression, and despite conservation efforts they could still be at risk (Shepard *et al.* 2005).

The native *O. apache* and *O. gilae* trout are similarly affected. *O. apache* are ESA listed as threatened and *O. gilae* trout as endangered due to habitat reductions and also hybridization and genetic introgression with introduced rainbow trout (Dowling & Childs 1992). It has been reported that about 65% of *O. apache* populations have some degree of introgression and one population is 100% introgressed (Rhymer & Simberloff 1996), and at least two populations of *O. gilae* (Iron Creek and Mc*K*enna Creek) have been lost due to hybridization (USFWS 2003). Therefore, it is of great concern to understand the extent of the introgression of hatchery fish into Mexican native trout in order to mitigate this effect.

Here, I focus on three main goals: 1) document the genetic biodiversity of native trout in northwestern Mexico, 2) infer the phylogeographic history of the Mexican trout complex, and 3) evaluate the extent of hybridization and genetic introgression from hatchery-raised fish into native trout; all these including samples from all the basins in

which native trout are known to occur.

3.2 Methods

3.2.1 *Tissue collection and DNA extraction*

Between 1994 and 2010 an exhaustive effort was made by the binational group "Truchas Mexicanas" to collect a total of 914 tissue samples ($\approx$1cm^2) from 42 localities (13 basins represented) in northwestern Mexico (Table 3.1; Figure 3.1). Also, 147 tissue samples from hatchery rainbow trout were obtained from four hatcheries located in different basins where native trout have been reported, and one hatchery located in Guachochi, Chihuahua that is presumably rearing *O. chrysogatser* (Table 3.1). Additionally, tissue samples from five West Fork Black River *O. apache* and five Gila River *O. gilae* were obtained.

Genomic DNA from 300 samples was extracted at the Centro de Investigaciones Biológicas del Noreste, S.C. (CIBNOR) in La Paz, Baja California, Mexico (for details see De los Santos-Camarillo 2008). DNA extraction of the remaining samples took place at the Southwest Fisheries Science Center in Santa Cruz, CA, USA. The samples were digested with proteinase *K*, followed by DNA extraction with a semi-automated membrane-based system (DNeasy 96 Tissue Kit) on a BioRobot 3000 (QIAGEN Inc.).

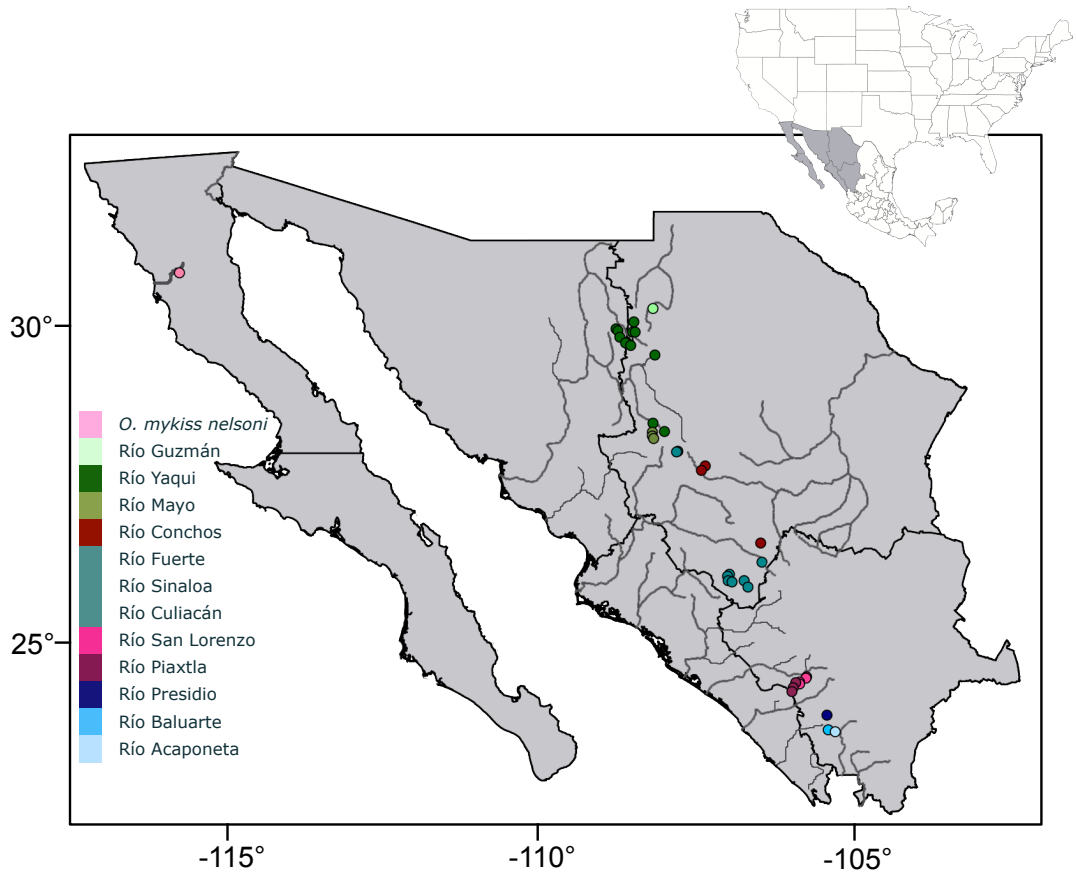

Figure 3.1: Geographic location of sampling sites from 13 major drainages in northwestern Mexico.

3.2.2 *Microsatellites and single nucleotide polymorphisms genotyping*

All individuals were genotyped at 18 microsatellite loci developed for salmonid species (Omy27 (McConnell *et al.* 1995a), Ssa289 (McConnell *et al.* 1995b); Omy77 (Morris *et al.* 1996); Ssa85 (O'Reilly *et al.* 1996); One11b, One13b (Scribner *et al.* 1996); Omy1011 (Condrey & Bentzen 1998); Ots103 (Small *et al.* 1998); Oki23 (Smith *et al.* 1998); Ots1b (Banks *et al.* 1999); OtsG3, OtsG43, OtsG85, OtsG243, OtsG249b, OtsG253, OtsG401, OtsG409 (Williamson *et al.* 2002)). This set of loci has proven to be highly informative in the study of *O. mykiss* population structure and interactions among different groups in California, including all major groups of hatchery rainbow trout (Aguilar & Garza 2006; Pearse *et al.* 2007; Clemento *et al.* 2009; Garza *et al.* 2014). Also, some of these markers have been tested in *O. clarkii* populations (Wenburg *et al.* 1998; Nielsen & Sage 2002) .

PCR was conducted using 4μL template DNA, 6.9μL H_2O, 1.5μL 10X PCR buffer (Applied Biosystems Inc.), 0.9μM $MgCl_2$, 0.6μM dNTPs, 1μM fluorescently labeled oligonucleotide primers, and 0.04U Amplitaq DNA polymerase (Applied Biosystems Inc.). PCR conditions consisted of 94°C for 3 min; then 9 cycles at 94°C for 30 s, 52 - 60°C for 2 min, and 72°C for 30 s; followed by 15 cycles at 92°C for 30 s, 52 - 60°C for 2 min, and 72°C for 30 s, with a final step at 72°C for 10 min. The PCR products were electrophoresed on an ABI377 genetic analyzer. Allele sizes were determined with Genotyper software (Applied Biosystems) and confirmed by two people independently.

A total of 93 single nucleotide polymorphism (SNP) locI was genotyped on all samples. These SNP markers include three loci from Aguilar and Garza (2008), six from Campbell *et al.* (2009), 82 from Abadía-Cardoso *et al.* (2011), and four unpublished. They have been validated in many populations from California, Oregon and Washington, as well as introduced populations in other parts of the world (A. A-C. & J.C.G., unpublished data). A PCR pre-amplification was carried out in 5.4μL aliquots containing 2.5μL of 2X Master Mix (QIAGEN Inc.), 1.3μM pooled oligonucleotide primers, and 1.6μL template DNA. Pre-amplification thermal cycling conditions included an initial denaturation of 15 min at 95°C, and 13 cycles of 15s at 95°C, 4 min at 60°C (+1°C/cycle). Pre-amplification PCR products were diluted 1:3 in 2 mM Tris. The genotyping method was the 5' nuclease allelic discrimination or TaqMan assay (Applied Biosystems) for high-throughput genotyping. The genotyping was carried out in 96.96 Dynamic SNP Genotyping Arrays on an EP1 System (Fluidigm Corporation) under the manufacturer's specifications.

Additionally, genotypes from the 18 microsatellite and 93 SNP loci from 18 natural-origin *O. mykiss* populations ($N = 675$) from California, USA that represent six Distinct Population Segments (DPSs) (Busby *et al.* 1996), and four *O. mykiss* hatchery strains ($N = 187$) were included in the analyses. Data from five cutthroat trout subspecies (coastal (*O. clarkii clarkii*; ($N = 47$)), Yellowstone (*O. clarkii bouvieri* ($N = 20$)), Bonneville (*O. clarkii utah* ($N = 16$)), Rio Grande (*O. clarkii virginalis* ($N = 10$)), and Colorado (*O. clarkii pleuriticus* ($N = 8$))) were also incorporated (Table 3.1). All these populations were carefully selected as the most representative lineages based

on previous studies (Bjorkstedt *et al.* 2005; Aguilar & Garza 2006; Clemento *et al.* 2009; Wilson & Turner 2009; Garza *et al.* 2014).

3.2.3 *Data analysis*

The two different classes of markers used have basic differences, such as a higher level of polymorphism in microsatellites, lower mutation rate in SNPs, different mutation process, among others. I believe that these differences could provide distinct and complementary information on the evolutionary history of these groups of trout. Therefore, some data analyses were performed for both SNPs and microsatellites separately.

Within population genetic variation was examined using different approaches. Expected (H_E) and observed (H_O) heterozygosities (Nei 1978) were estimated using GENEPOP (Rousset 2008) for microsatellites and SNPs separately. Percentage of polymorphic SNPs (P) at 0.95 and 0.99 was calculated using GENETIX 4.05 (Belkhir *et al.*). I used the package hierfstat for R (Goudet 2005; R Development Core Team 2011) to estimate microsatellite allelic richness by rarefaction (A_R) to correct for sample size differences. A Bayesian analysis of group determination implemented in the program STRUCTURE 2.2 (Pritchard *et al.* 2000) was performed. This analysis, based on individual multilocus genotypes with no prior geographic information of the populations, indicates the level of mixing within and between groups. Values of K = 2 - 7 were used, and 20 iterations were executed for each value of K with a burn-in period of 50,000 steps and 150,000 Monte Carlo Markov Chain replicates. The results from

these STRUCTURE runs were reordered and visualized using the software CLUMPP (Jakobsson & Rosenberg 2007) and Distruct (Rosenberg 2004).

Relationships among populations were explored using three methods: a) Pairwise F_{ST} tests and their significance levels through 10,000 permutations were estimated using Arlequin 3.5 (Excoffier & Lischer 2010); b) Principal Components Analysis (PCA) was performed using the R based package adegenet 1.3-4 (Jombart 2008); and c) Unrooted phylogeographic neighbor-joining trees were created using PHYLIP (Felsenstein 2005) for the microsatellite data only and with the combined dataset. Markers that failed for an entire population were excluded from the PHYLIP analysis, leaving 12 microsatellite and 85 SNP loci. I used the Cavalli-Sforza and Edwards (1967) method to estimate pairwise genetic distances and 1,000 bootstrapped distance matrices to evaluate node support. I excluded from neighbor-joining trees those populations with less than eight individuals.

Table 3.1: Populations used in this study from north to south. N: total number of samples; Lat: Latitude; Long: Longitude; *: estimated location; DPS: Klamath Mountains Province ([a]); Northern California ([b]); Central Valley ([c]); Central California Coast ([d]); South-Central California Coast ([e]); Southern California ([f]).

Pop. num.	Population name	N	Lat	Long	Collection year	Species
1	Klamath River-Blue Creek	32	41.44	-123.91	2001	*O. mykiss*[a]
2	Mattole River-South Fork Bear Creek	31	40.03	-124.02	2001	*O. mykiss*[b]
3	Gualala River-Fuller Creek	29	38.70	-123.33	2001	*O. mykiss*[b]
4	Sacramento River-McCloud River-Claiborne Creek	33	41.06	-122.12	2006	*O. mykiss*[c]
5	Sacramento River-Battle Creek	47	40.38	-122.14	2005	*O. mykiss*
6	Sacramento River-Deer Creek	45	40.00	-121.97	2003	*O. mykiss*
7	Sacramento River-Upper Yuba River	26	39.57	-120.72	2006	*O. mykiss*
8	Sacramento River-North Fork American River	49	39.20	-120.61	2006	*O. mykiss*[c]
9	SanFrancisquito Creek-Los Trancos Creek	24	37.41	-122.19	2001	*O. mykiss*[d]
10	Waddell Creek	31	37.12	-122.27	2001	*O. mykiss*[d]
11	Carmel River	32	36.41	-121.68	2001	*O. mykiss*[e]
12	Chorro Creek-Pennington Creek	31	35.34	-120.73	2001	*O. mykiss*[e]
13	SantaMaria River-Sisquoc River	47	34.88	-120.29	2011	*O. mykiss*[f]
14	Santa Ynez River-Salsipuedes Creek	47	34.62	-120.39	2003	*O. mykiss*[f]
15	Ventura Creek-North Fork Matilija Creek	47	34.51	-119.38	2003	*O. mykiss*[f]
16	Santa Clara River-Sespe Creek-Lion Canyon	47	34.38	-118.96	2003	*O. mykiss*[f]
17	Pauma Creek	47	33.34	-116.97	1997	*O. mykiss*[f]
18	Sweetwater River	37	32.92	-116.57	1997	*O. mykiss*[f]
19	Río Santo Domingo-Arroyo San Antonio and Arroyo La Grulla	42	30.82*	-115.63	1994	*O. mykiss nelsoni*
20	Río Guzmán-Río Piedras Verdes-Arroyo Escalariado	27	30.26	-108.20	2005	*O. sp.*
21	Río Yaqui-Río Bavispe-Arroyo Las Guacamayas	23	30.05	-108.50	2007	*O. sp.*
22	Río Yaqui-Río Bavispe-Arroyo La Nutria	23	29.94	-108.78	2007	*O. sp.*
23	Río Yaqui-Río Bavispe-Arroyo Los Cuarteles	26	29.91	-108.75	2007	*O. sp.*
24	Río Yaqui-Río Bavispe-Río Gavilán-Arroyo Yenquin	20	29.89	-108.52	2009	*O. sp.*
25	Río Yaqui-Río Bavispe-Río Gavilán-Arroyo Las Truchas	20	29.89	-108.48	2009	*O. sp.*
26	Río Yaqui-Río Bavispe-Arroyo La Presita	12	29.81	-108.72	2005	*O. sp.*
27	Río Yaqui-Río Bavispe-Arroyo El Arco-Arroyo Pedernal	10	29.73	-108.61	2007	*O. sp.*
28	Río Yaqui-Río Bavispe-Arroyo El Arco-Arroyo Largo	20	29.72	-108.63	2007	*O. sp.*
29	Río Yaqui-Río Bavispe-Río Negro-Arroyo El Cocoño	16	29.68	-108.55	2007	*O. sp.*

Table 3.1 Continued

Pop. num.	Population name	N	Lat	Long	Collection year	Species
30	Río Yaqui-Río Sirupa-Río Papagochi-Arroyo El Salto	21	29.53	-108.17	2007	*O. sp.*
31	Río Yaqui-Río Sirupa-Río Papagochi-Río Tutuaca	13	28.46	-108.20	2005	*O. sp.*
32	Río Yaqui-Río Sirupa-Arroyo Banderella	15	28.33	-108.02	2005	*O. sp.*
33	Río Mayo-Arroyo Concheño	36	28.32	-108.21	2009	*O. sp.*
34	Río Mayo-Río Candameña	15	28.26	-108.21	2005	*O. sp.*
35	Río Mayo-Arroyo La Estrella	19	28.22	-108.19	2009	*O. sp.*
36	Río Conchos-Arroyo Ureyna	16	27.79	-107.38	2006	*O. sp.*
37	Río Conchos-Río Rituchi	9	27.72	-107.44	2005	*O. sp.*
38	Río Conchos-Río El Porvenir-Arroyo San Antonio-Arroyo El Molino	28	26.58	-106.51	2007	*O. sp.*
39	Río Fuerte-Río Ateros-Arroyo Aparique	8	28.02	-107.81	2005	*O. chrysogaster*
40	Río Fuerte-Arroyo San Vicente	10	28.01	-107.83	2009	*O. chrysogaster*
41	Río Fuerte-Río Verde	30	26.28	-106.49	2007	*O. chrysogaster*
42	Río Fuerte-Arroyo Santa Rosalia-Arroyo Las Truchas	29	26.09	-107.00	2008	*O. chrysogaster*
43	Río Sinaloa-Río San José-Río Basonopira-Arroyo Potrero	28	26.06	-107.03	2008	*O. chrysogaster*
44	Río Sinaloa-Arroyo Rancho de El Medio-below waterfall	10	25.99	-106.77	2007	*O. chrysogaster*
45	Río Sinaloa-Arroyo Rancho de El Medio-above waterfall	10	25.99	-106.77	2007	*O. chrysogaster*
46	Río Sinaloa-Río Mohinora-Arroyo Soldado	40	25.99	-107.02	2008	*O. chrysogaster*
47	Río Sinaloa-Arroyo Hondo	21	25.97	-106.96	2008	*O. chrysogaster*
48	Río Culiacán-Arroyo Santa Rosa	21	25.89	-106.71	2007	*O. chrysogaster*
49	Río San Lorenzo-Río Los Remedios-Arroyo La Sidra above waterfall	16	24.48	-105.79	2004	*O. sp.*
50	Río San Lorenzo-Río Los Remedios-Arroyo La Sidra below waterfall	19	24.46	-105.80	2004	*O. sp.*
51	Río San Lorenzo-Río Los Remedios-Arroyo Las Truchas	28	24.40	-105.90	2004	*O. sp.*
52	Río San Lorenzo-Río Los Remedios-Arroyo Las Veredas	40	24.37	-105.90	2004	*O. sp.*
53	Río Piaxtla-Río Verde-Arroyo SanAntonio	11	24.39	-105.96	2004	*O. sp.*
54	Río Piaxtla-Arroyo del Granizo	26	24.31	-106.00	2004	*O. sp.*
55	Río Piaxtla-Arroyo de la Plazuela-ejido El Maguey	34	24.31	-106.07	2010	*O. sp.*
56	Río Piaxtla-Arroyo Palo Berdal	28	24.29	-106.05	2010	*O. sp.*
57	Río Piaxtla-Arroyo Cruz Larga	25	24.25	-106.02	2004	*O. sp.*
58	Río Presidio-Arroyo Quebrada de la Vega	15	23.88	-105.47	2004	*O. sp.*
59	Río Baluarte-Arroyo Santa Barbara	37	23.65	-105.45	2004	*O. sp.*
60	Río Acaponeta-Arroyo Las Cebollas	17	23.62	-105.34	2004	*O. sp.*
61	West Fork Black River	5	33.88	-109.47	2000	*O. apache*
62	Gila River-Main Diamond Creek	5	N/A	N/A	2004	*O. gilae*

Table 3.1 Continued

Pop. num.	Population name	N	Lat	Long	Collection year	Species
63	Maple Creek	47	N/A	N/A	2002	*O. clarkii clarkii*
64	Snake River-Barnes Creek	20	N/A	N/A	N/A	*O. clarkii bouvieri*
65	Bonneville-Glenwood	16	N/A	N/A	N/A	*O. clarkii utah*
66	Colorado River	8	N/A	N/A	N/A	*O. clarkii pleuriticus*
67	Rio Grande	10	N/A	N/A	N/A	*O. clarkii virginalis*
68	Río Yaqui-Río Bavispe-Hatchery Arroyo Yenquin	24	29.89	-108.49	2009	*O. mykiss*
69	Río Yaqui-Río Bavispe-Arroyo El Arco-Hatchery Truchas la Presita	20	29.73	-108.62	2007	*O. mykiss*
70	Río Fuerte-Río Ateros-Arroyo Aparique abandoned hatchery	7	28.02	-107.81	2005	*O. chrysogaster*
71	Río SanLorenzo-Río Los Remedios-Hatchery Piscicultura Vencedores	18	24.47	-105.79	2004	*O. mykiss*
72	Hatchery Centro trutícola Guachochi	78				*O. chrysogaster*
73	Sacramento River-American River-Coleman Strain	46	38.63	-121.23	2003	*O. mykiss*
74	Owens Lake Basin-Hot Creek-Kamloops Strain	47	37.65	-118.84	2005	*O. mykiss*
75	Sacramento River-American River-Eagle Lake Strain	47	38.63	-121.23	2005	*O. mykiss*
76	Sacramento River-American River-Mt. Shasta Strain	47	38.63	-121.23	2002	*O. mykiss*

3.2.4 *Genetic introgression*

Due to evident establishment of non-native hatchery rainbow trout in many, if not all, of the main drainages of northwestern Mexico (García de León, pers. comm.), I explored the potential genetic introgression from these hatchery trout into the native trout populations using two different approaches: a Bayesian analysis with STRUCTURE ($K = 2$ - 5, five iterations each) and PCA. These analyses were performed using all the natural-origin populations from a basin (*e.g.* Río San Lorenzo) and fish from hatcheries established in that basin (*e.g.* "Piscicultura Vencedores" hatchery), along with fish from California hatchery strains. Even though I do not have samples from hatcheries located in the southernmost basins, I explored the genetic introgression from California hatchery rainbow trout in ríos Presidio, Baluarte, and Acaponeta, based on the results from previous analysis.

3.3 Results

I successfully genotyped 1,055 trout from northwestern Mexico, *O. apache*, *O. gilae*, and Mexican hatcheries with the microsatellite panel, and 1,027 with the SNP panel. I excluded from the analysis those individuals that had excessive missing data ($\geq$ 10 missing SNP loci and $\geq$ 9 missing microsatellite loci). A total of 1,999 fish from Mexican and California populations were included in the final microsatellite analyses and 1,985 in the SNP analyses (Table 3.2).

Observed heterozygosity per population with microsatellites ranged from 0.033

in Río Conchos-Arroyo Ureyna to 0.736 in Klamath River-Blue Creek; and ranged from zero in several populations from NSMO and *O. clarkii* ssp. to 0.413 in Gualala River-Fuller Creek with SNPs (Table 3.2; Figure 3.2). Overall, heterozygosity was higher for both marker types in *O. mykiss* populations (microsatellites: mean (H_O) = 0.636, range = 0.334 - 0.736; SNPs: mean (H_O) = 0.336, range = 0.129 - 0.413), than in any of the SMO groups (microsatellites: NSMO mean (H_O) = 0.288, range = 0.033 - 0.516; *O. chrysogaster* mean (H_O) = 0.354, range = 0.190 - 0.606; SSMO mean (H_O) = 0.382, range = 0.191 - 0.624; SNPs: NSMO mean (H_O) = 0.008, range = 0 - 0.040; *O. chrysogaster* mean (H_O) = 0.074, range = 0.003 - 0.113; SSMO mean (H_O) = 0.144, range = 0.016 - 0.324; Figure 3.2), and than the other species (microsatellites: *O. apache* (H_O) = 0.44; *O. gilae*: (H_O) = 0.192; *O. clarkii* ssp. mean (H_O) = 0.379, range = 0.281 - 0.537; SNPs: *O. apache* (H_O) = 0.006; *O. gilae*: (H_O) = 0.002; *O. clarkii* ssp. mean (H_O) = 0.008, range = 0 - 0.031; Table 3.2; Figure 3.2).

Mean number of alleles per microsatellite and A_R were highest in *O. mykiss* populations (mean alleles/locus = 7.74; mean A_R = 1.65). Río Conchos-Río Rituchi showed the lowest number of alleles per microsatellite and A_R (alleles/locus = 1.06; A_R = 1.03; Table 3.2; Figure 3.3). Within Mexican trout, the highest number of alleles per microsatellite and highest A_R were observed in Río Fuerte-Río Verde (alleles/locus = 8.0; A_R = 1.65; Table 3.2).

Percentage of polymorphic SNP loci at $P(0.95)$ ranged from 0 to 0.99 and at $P(0.99)$ from 0 to 1, with the highest values observed again in *O. mykiss* (mean $P(0.95)$ = 0.88; mean $P(0.99)$ = 0.95). The lowest values of P were observed in NSMO

(mean $P(0.95) = 0.02$; mean $P(0.99) = 0.05$), where all locI was monomorphic in several populations (Table 3.2; Figure 3.3).

I observed higher mean F_{ST} values with SNPs than microsatellites (Table 3.2). Significant negative correlations were found when F_{ST} values were compared to both microsatellite allelic richness (A_R: $F_{1,74} = 517.8$, $R^2 = 0.875$, $p < 0.001$) and percentage of polymorphic SNPs ($P(0.95)$: $F_{1,74} = 859.5$, $R^2 = 0.921$, $p < 0.001$; $P(0.99)$: $F_{1,74} = 591.6$, $R^2 = 0.889$, $p < 0.001$; Figure 3.3) as expected, since it has been demonstrated that the level of heterozygosity directly affects the level of differentiation among groups (Hedrick 1999).

The STRUCTURE analysis clustered the individuals according to geographic location for *O. mykiss* and Mexican populations. However, contrary to my expectations, I did not see a pattern of clustering for the different species. *O. clarkii* ssp. shared ancestry with *O. chrysogaster*, while *O. gilae*, and *O. apache* present shared ancestry with *O. mykiss*, *O. chrysogaster*, and also with NSMO (Figure 3.4).

Table 3.2: Summary statistics. N: number of individuals succsesfully genotyped with the microsatellite panel and the SNP panel respectively; Expected (H_E) and observed (H_O) heterozygosity; allelic richness (A_R); percentage of polymirphic loci (P) at 0.95 and 0.99.

		Msats						SNPs						
Group	Population	N	Num. of loci	H_E	H_O	Alleles/ locus	A_R	Mean pairwise F_{ST}	N	Num. of loci	H_E	H_O	P (0.95)	P (0.99)	Mean pairwise F_{ST}
O. m. nelsoni	Klamath River-Blue Creek	32	18	0.75	0.74	11.56	1.75	0.20	32	92	0.33	0.31	0.90	0.97	0.35
	Mattole River-South Fork Bear Creek	31	17	0.70	0.73	8.00	1.70	0.24	31	92	0.35	0.35	0.88	0.97	0.35
	Gualala River-Fuller Creek	29	17	0.68	0.68	8.47	1.68	0.24	29	92	0.40	0.41	0.99	0.99	0.30
	Sacramento River-Claiborne Creek	30	18	0.66	0.64	7.22	1.66	0.26	30	93	0.32	0.32	0.83	0.92	0.37
	Sacramento River-Battle Creek	47	17	0.68	0.66	12.24	1.68	0.21	47	93	0.37	0.35	0.99	1.00	0.29
	Sacramento River-Deer Creek	43	18	0.74	0.72	12.89	1.74	0.22	43	93	0.37	0.38	0.98	1.00	0.29
	Sacramento River-Upper Yuba River	26	18	0.72	0.71	8.33	1.72	0.23	26	93	0.38	0.39	0.96	1.00	0.32
	Sacramento River-North Fork American River	48	18	0.67	0.65	8.72	1.67	0.24	48	93	0.33	0.34	0.90	0.98	0.35
	SanFrancisquito Creek-Los Trancos Creek	24	17	0.62	0.65	6.53	1.62	0.28	24	92	0.36	0.37	0.89	0.92	0.38
	Waddell Creek	31	18	0.63	0.62	7.22	1.63	0.25	31	92	0.37	0.37	0.95	0.96	0.36
	Carmel River	32	18	0.67	0.68	8.89	1.67	0.24	32	92	0.38	0.37	0.97	0.98	0.35
	Chorro Creek-Pennington Creek	31	18	0.63	0.64	5.56	1.63	0.28	31	92	0.33	0.35	0.83	0.90	0.40
	SantaMaria River-Sisquoc River	47	18	0.63	0.63	6.28	1.63	0.30	47	93	0.32	0.32	0.90	0.96	0.37
	Santa Ynez River-Salsipuedes Creek	47	18	0.61	0.57	6.50	1.61	0.25	47	93	0.33	0.32	0.89	0.95	0.36
	Ventura Creek-North Fork Matilija Creek	46	18	0.62	0.62	5.72	1.62	0.27	46	93	0.33	0.33	0.86	0.96	0.37
	Santa Clara River-Sespe Creek-Lion Canyon	47	18	0.60	0.60	7.94	1.60	0.30	47	93	0.31	0.31	0.87	0.98	0.38
	Pauma Creek	47	14	0.64	0.64	6.29	1.64	0.24	43	93	0.34	0.34	0.86	0.97	0.33
	Sweetwater River	37	14	0.60	0.58	5.36	1.60	0.27	37	93	0.33	0.33	0.86	0.94	0.37
	Río Santo Domingo	40	18	0.41	0.33	3.28	1.41	0.35	39	93	0.15	0.13	0.41	0.54	0.61
NSMO	Río Guzmán-Arroyo Escalariado	27	18	0.13	0.12	1.61	1.13	0.62	26	93	0.01	0.01	0.01	0.18	0.68
	Río Yaqui-Bavispe-Arroyo Las Guacamayas	23	18	0.35	0.31	2.56	1.35	0.49	22	92	0.01	0.00	0.02	0.03	0.66
	Río Yaqui-Bavispe-Arroyo La Nutria	23	18	0.41	0.39	2.72	1.41	0.42	23	92	0.01	0.01	0.02	0.03	0.64
	Río Yaqui-Bavispe-Arroyo Los Cuarteles	26	17	0.27	0.25	1.94	1.27	0.46	25	92	0.00	0.00	0.00	0.00	0.74
	Río Yaqui-Bavispe-Arroyo Yenquin	20	18	0.48	0.48	2.94	1.48	0.38	20	90	0.01	0.01	0.03	0.04	0.62
	Río Yaqui-Bavispe-Arroyo Las Truchas	17	18	0.16	0.16	1.61	1.16	0.51	18	90	0.01	0.00	0.02	0.02	0.64
	Río Yaqui-Bavispe-Arroyo La Presita	12	18	0.45	0.39	2.89	1.45	0.38	12	92	0.01	0.01	0.03	0.03	0.61
	Río Yaqui-Bavispe-Arroyo Pedernal	10	18	0.47	0.46	2.89	1.47	0.36	10	92	0.01	0.01	0.03	0.03	0.60
	Río Yaqui-Bavispe-Arroyo Largo	16	18	0.47	0.45	3.61	1.47	0.34	14	92	0.04	0.04	0.08	0.34	0.56
	Río Yaqui-Bavispe-Arroyo El Cocoño	16	18	0.41	0.39	3.06	1.41	0.36	16	91	0.01	0.02	0.04	0.04	0.62
	Río Yaqui-Sirupa-Arroyo El Salto	21	18	0.18	0.19	1.89	1.18	0.50	20	92	0.01	0.02	0.04	0.12	0.65
	Río Yaqui-Sirupa-Río Tutuaca	13	18	0.42	0.38	2.78	1.42	0.42	13	92	0.01	0.00	0.02	0.02	0.65
	Río Yaqui-Sirupa-Arroyo Banderella	15	18	0.50	0.52	3.44	1.50	0.37	15	92	0.00	0.00	0.00	0.00	0.68
	Río Mayo-Arroyo Concheño	36	17	0.19	0.19	1.94	1.19	0.54	36	91	0.00	0.00	0.00	0.00	0.76
	Río Mayo-Río Candameña	15	18	0.29	0.28	2.33	1.29	0.48	15	92	0.01	0.01	0.02	0.02	0.66
	Río Mayo-Arroyo La Estrella	19	18	0.26	0.27	2.06	1.26	0.46	19	91	0.01	0.02	0.03	0.03	0.66
	Río Conchos-Arroyo Ureyna	15	18	0.03	0.03	1.17	1.03	0.64	15	91	0.00	0.00	0.00	0.00	0.68
	Río Conchos-Río Rituchi	9	18	0.03	0.06	1.06	1.03	0.62	8	90	0.01	0.00	0.01	0.01	0.65
	Río Conchos-Arroyo El Molino	25	18	0.26	0.15	2.44	1.26	0.58	21	91	0.00	0.00	0.00	0.00	0.76
O. chrysogaster	Río Fuerte-Arroyo Aparique	8	18	0.22	0.19	2.17	1.22	0.44	8	93	0.05	0.05	0.37	0.37	0.69
	Río Fuerte-Arroyo San Vicente	10	18	0.27	0.27	2.28	1.27	0.47	10	93	0.11	0.10	0.47	0.47	0.64
	Río Fuerte-Río Verde	30	18	0.64	0.61	8.00	1.64	0.36	30	93	0.02	0.02	0.06	0.10	0.74
	Río Fuerte-Arroyo Las Truchas	29	18	0.35	0.36	3.00	1.35	0.44	29	91	0.06	0.06	0.16	0.20	0.68
	Río Sinaloa-Arroyo Potrero	28	18	0.25	0.24	2.72	1.25	0.54	26	91	0.01	0.00	0.02	0.02	0.77
	Río Sinaloa-Arroyo El Medio-below waterfall	10	18	0.44	0.39	2.78	1.44	0.42	10	93	0.11	0.11	0.30	0.30	0.61
	Río Sinaloa-Arroyo El Medio-above waterfall	10	18	0.37	0.33	2.44	1.37	0.47	9	93	0.10	0.09	0.26	0.26	0.63
	Río Sinaloa-Arroyo Soldado	40	18	0.35	0.36	4.00	1.35	0.36	38	91	0.10	0.10	0.26	0.37	0.60
	Río Sinaloa-Arroyo Hondo	21	18	0.41	0.36	3.61	1.41	0.36	21	92	0.09	0.09	0.25	0.29	0.60
	Río Culiacán-Arroyo Santa Rosa	21	18	0.51	0.43	3.28	1.51	0.30	21	93	0.11	0.10	0.30	0.34	0.60
SSMO	Río San Lorenzo-Arroyo La Sidra above waterfall	14	18	0.48	0.50	3.83	1.48	0.34	14	93	0.25	0.24	0.66	0.73	0.45
	Río San Lorenzo-Arroyo La Sidra below waterfall	17	18	0.59	0.60	4.78	1.59	0.33	17	93	0.31	0.32	0.85	0.88	0.40

Table 3.2 Continued

		Msats							SNPs					
Group	Population	N	Num. of loci	H_E	H_O	Alleles/ locus	A_R	Mean pairwise F_{ST}	N	Num. of loci	H_E	H_O	P (0.95)	P (0.99)	Mean pairwise F_{ST}
	Río San Lorenzo-Arroyo Las Truchas	28	18	0.26	0.26	3.61	1.26	0.45	27	93	0.08	0.08	0.34	0.67	0.63
	Río San Lorenzo-Arroyo Las Veredas	40	18	0.21	0.19	3.11	1.21	0.52	40	93	0.04	0.04	0.16	0.58	0.70
	Río Piaxtla-Arroyo SanAntonio	11	18	0.26	0.30	2.11	1.26	0.53	11	93	0.07	0.07	0.18	0.31	0.63
	Río Piaxtla-Arroyo del Granizo	26	18	0.20	0.19	1.67	1.20	0.46	26	93	0.02	0.02	0.04	0.04	0.72
	Río Piaxtla-Arroyo de la Plazuela	34	18	0.30	0.31	2.83	1.30	0.45	33	93	0.04	0.04	0.11	0.11	0.69
	Río Piaxtla-Arroyo Palo Berdal	28	18	0.27	0.27	2.22	1.27	0.47	27	93	0.03	0.03	0.09	0.10	0.69
	Río Piaxtla-Arroyo Cruz Larga	25	18	0.23	0.25	1.67	1.23	0.53	25	93	0.03	0.03	0.08	0.08	0.70
	Río Presidio-Arroyo Quebrada de la Vega	15	18	0.64	0.62	5.17	1.64	0.27	15	93	0.30	0.29	0.80	0.84	0.42
	Río Baluarte-Arroyo Santa Barbara	37	18	0.51	0.47	4.67	1.51	0.35	35	93	0.24	0.22	0.71	0.83	0.46
	Río Acaponeta-Arroyo Las Cebollas	17	18	0.59	0.62	4.28	1.59	0.29	17	93	0.29	0.32	0.81	0.83	0.43
O. apache	West Fork Black River	5	14	0.46	0.45	2.43	1.46	0.40	5	93	0.01	0.01	0.01	0.01	0.70
O. gilae	Gila River-Main Diamond Creek	5	13	0.20	0.19	1.69	1.20	0.46	5	91	0.01	0.00	0.01	0.01	0.71
O. clarkii ssp.	Maple Creek	47	17	0.62	0.54	7.47	1.62	0.40	47	91	0.04	0.03	0.15	0.27	0.71
	Snake River-Barnes Creek	9	13	0.42	0.37	4.15	1.42	0.40	19	90	0.01	0.01	0.02	0.02	0.73
	Bonneville-Glenwood	16	13	0.49	0.42	3.31	1.49	0.48	16	90	0.00	0.00	0.00	0.00	0.73
	Colorado River	4	14	0.31	0.29	2.14	1.31	0.57	8	89	0.00	0.00	0.00	0.00	0.71
	Rio Grande	10	14	0.45	0.28	3.43	1.45	0.47	10	90	0.00	0.00	0.00	0.00	0.73
Mexican	Río Yaqui-Bavispe-Hatchery Arroyo Yenquin	24	18	0.70	0.69	6.94	1.70	0.23	24	93	0.34	0.33	0.88	0.96	0.36
hatcheries	Río Yaqui-Bavispe-Hatchery Truchas la Presita	20	18	0.71	0.73	6.17	1.71	0.24	19	93	0.33	0.34	0.90	0.94	0.39
	Río Fuerte-Arroyo Aparique abandoned hatchery	7	18	0.60	0.65	3.78	1.60	0.31	7	93	0.27	0.28	0.74	0.74	0.50
	Río SanLorenzo-Hatchery Piscicultura Vencedores	18	18	0.69	0.68	6.94	1.69	0.23	18	93	0.36	0.36	0.94	0.97	0.34
	Hatchery Centro trutcola Guachochi	75	18	0.62	0.61	5.94	1.62	0.24	73	93	0.33	0.36	0.91	0.92	0.33
California	Coleman Strain	46	18	0.62	0.61	6.44	1.62	0.26	46	93	0.34	0.34	0.90	0.95	0.32
hatcheries	Kamloops Strain	47	17	0.61	0.59	7.47	1.61	0.26	47	92	0.23	0.23	0.73	0.80	0.44
	Eagle Lake Strain	47	17	0.60	0.59	5.18	1.60	0.29	47	93	0.25	0.24	0.75	0.94	0.40
	Mt. Shasta Strain	47	17	0.60	0.56	5.24	1.60	0.28	47	93	0.31	0.31	0.80	0.88	0.38

Some general patterns can be observed across the different K values. For example, clear breaks occur between *O. mykiss*, NSMO, *O. chrysogaster*, and ríos San Lorenzo and Piaxtla. Within the Río Conchos populations I observed an unexpected pattern. On one hand, ríos Rituchi and Ureyna cluster with the rest of the NSMO complex, and on the other hand, Arroyo El Molino shares ancestry with both the NSMO complex and *O. chrysogaster*. I saw the same situation for Río Fuerte-Río Verde. Within *O. chrysogaster* I observed two populations that do not follow the same pattern as others. Río Fuerte-Arroyo Aparique and Río Fuerte-Arroyo San Vicente clearly share ancestry with *O. mykiss* -maybe an indication of genetic introgression. The Ríos Presidio, Baluarte and Acaponeta also show admixture with *O. mykiss* and more specifically with the Central Valley DPS populations and hatchery strains (Figure 3.4).

Highly significant genetic differentiation was documented based on the estimated pairwise F_{ST} values when both microsatellites and SNPs were combined (Tables S2 and S3). The strongest differentiation was observed among SMO populations (mean pairwise F_{ST} = 0.39 - 0.73) and between SMO and *O. mykiss* populations.

The PCA revealed seven well-differentiated clusters (Figure 3.5). The first cluster (dark green) corresponds to the NSMO, and comprises all the localities from both Río Yaqui tributaries (ríos Bavispe and Sirupa), Río Guzmán, Río Mayo, and two tributaries from the Río Conchos (ríos Rituchi and Ureyna). A second cluster (dark pink) includes all tributaries from ríos Piaxtla and San Lorenzo except for Arroyo La Sidra (above and below waterfall; see below for more information about this locality). A third cluster (yellow/orange) encompasses all *O. mykiss* populations, including *O. m.*

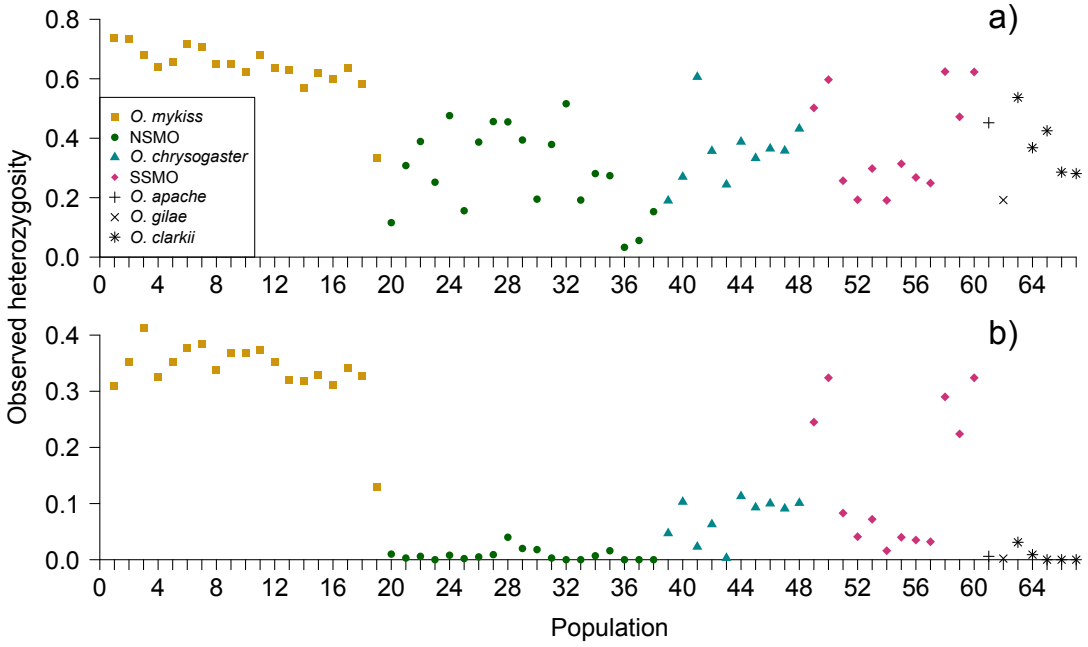

Figure 3.2: Mean observed heterozygosity per population. a) Microsatellites; b) SNPs. The populations are organized from north to south except for populations 61 to 67, which correspond to *O. apache*, *O. gilae*, and *O. clarkii*. NSMO: Northern Sierra Madre Occidental; SSMO: Southern Sierra Madre Occidental.

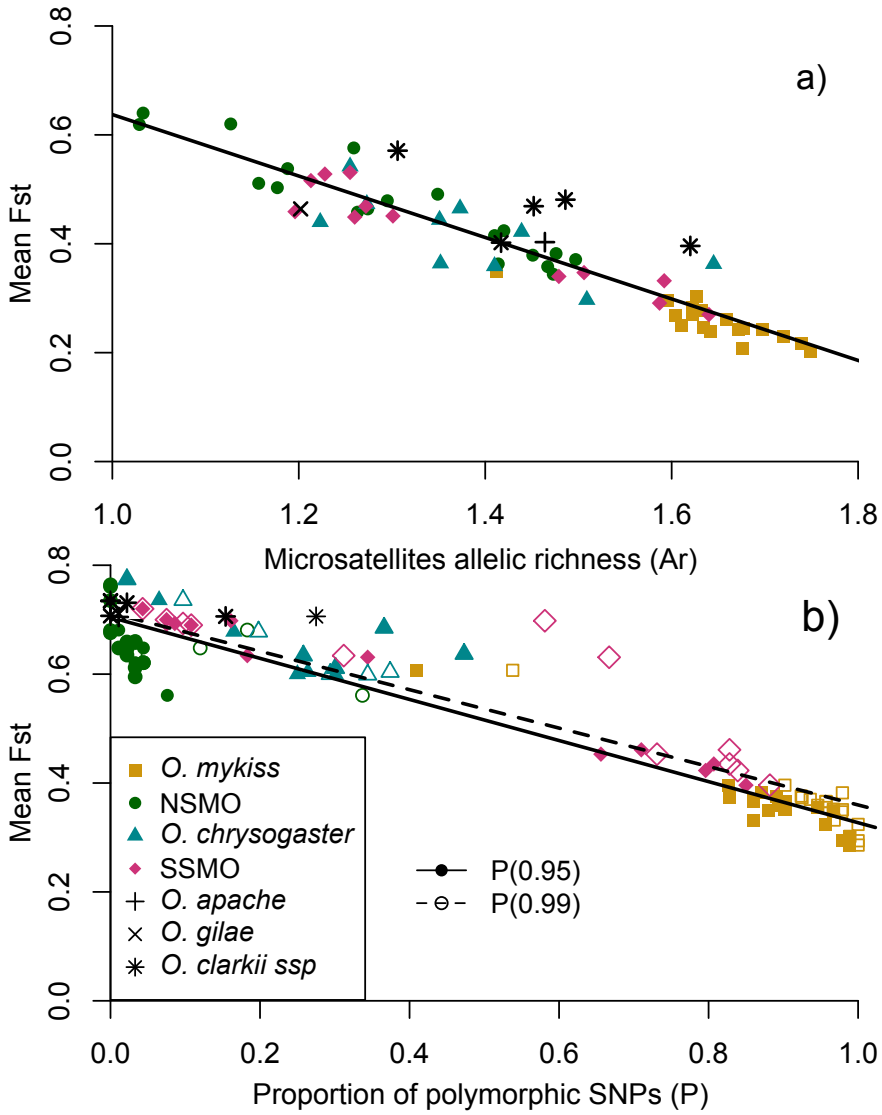

Figure 3.3: Correlation between mean F_{ST} and genetic diversity. a) Microsatellite allelic richness and b) percentage of polymorphic SNPs at 0.95 and 0.99. Populations are organized from north to south except for populations 61 to 67, which correspond to *O. apache*, *O. gilae*, and *O. clarkii*. NSMO: Northern Sierra Madre Occidental; SSMO: Southern Sierra Madre Occidental.

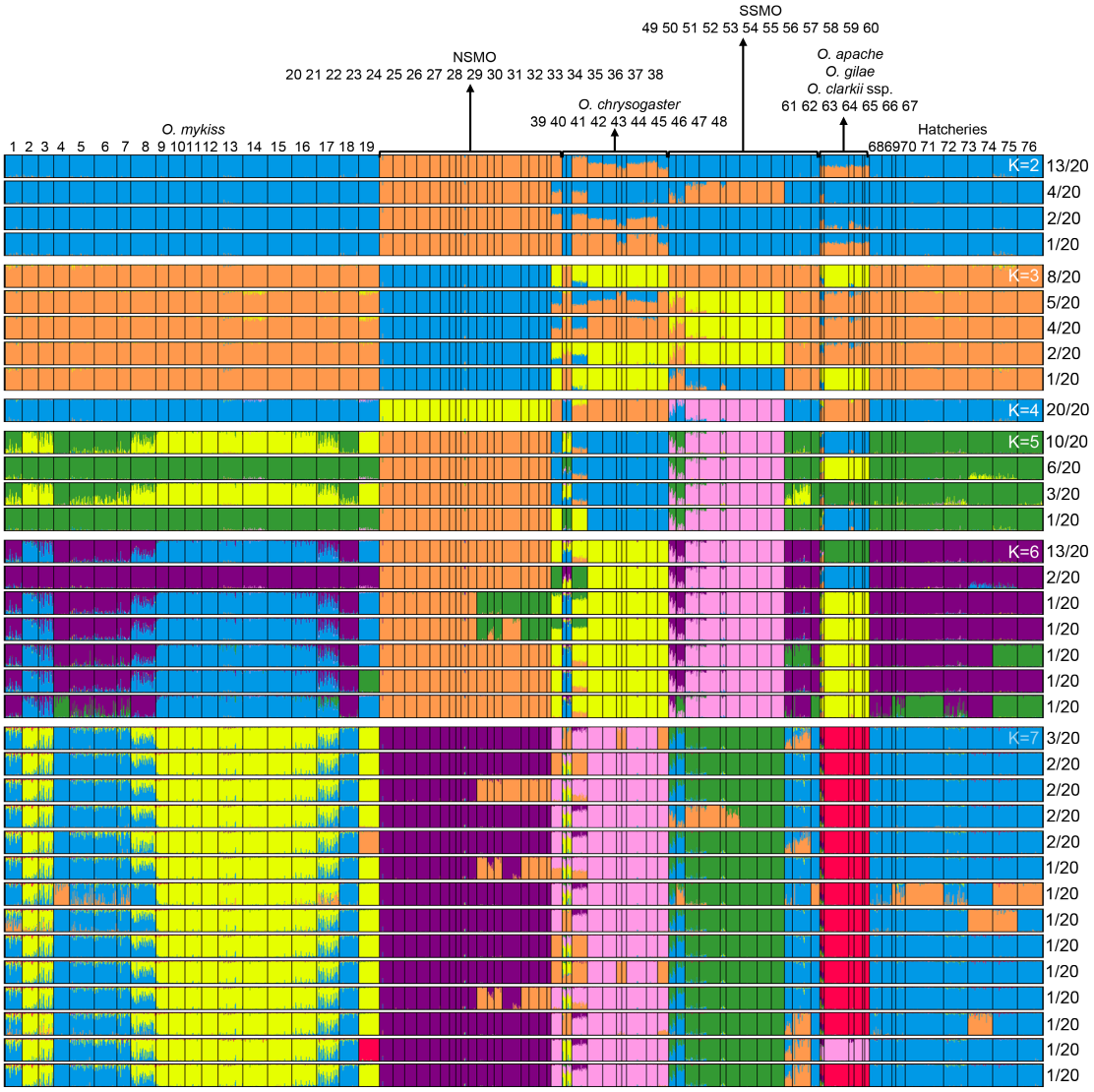

Figure 3.4: STRUCTURE analysis. Estimated membership fraction (represented by color proportions) of 1,932 individuals from 18 *O. mykiss* populations, 42 trout populations from northwestern Mexico, one *O. gilae* and one *O. apache* population, five *O. clarkii* subspecies, and five Mexican and four California rainbow trout stocks, using 18 microsatellites and 93 SNPs. Horizontal plots represent STRUCTURE runs constructed with Distruct. Each thin, colored, vertical line represents one individual. Vertical black lines separate collection localities. A summary of the 20 runs for each K value ($K = 2$ - 7) is shown. The right column indicates the number of observations for that specific pattern. NSMO: Northern Sierra Madre Occidental; SSMO: Southern Sierra Madre Occidental. Numbers on top represent the Population number in table S1.

nelsoni, all California and Mexican hatcheries, as well as fishes from the SSMO southernmost localities (ríos Presidio, Baluarte and Acaponeta). A fourth cluster (green/blue) includes two localities: Río Conchos-Arroyo El Molino and Río Fuerte-Río Verde, the last being previously described as *O. chrysogaster*. The rest of the *O. chrysogaster* localities form a fifth group (light blue), except for two tributaries from Río Fuerte (arroyos Aparique and San Vicente). These Río Fuerte tributaries surprisingly cluster with the *O. mykiss* group when PC1 and PC2 are plotted (Figure 3.5a) and with the San Lorenzo/Piaxtla cluster when PC1 and PC3 are used. *O. gilae* and *O. apache* trout define the sixth group (olive green). Finally, *O. clarkii* subspecies (light green) overlap with two *O. chrysogaster* populations (Río Sinaloa-Arroyo El Potrero and Río Fuerte-Arroyo Las Truchas) when PC1 and PC2 are plotted (Figure 3.5a) but separate when PC1 and PC3 are used (Figure 3.5b).

Overall, topologies were concordant between the two unrooted phylogenetic trees (Figure 3.6 and Figure 3.7), with the exception of the southern populations ríos Presidio and Baluarte that cluster within the *O. mykiss* lineage on the combined tree while they form a separate group on the microsatellite tree.

Several noticeable features can be identified in the population grouping patterns in both trees. First, the topology observed is mostly consistent with the different species as well as with the geographic proximity of streams; clustering all the *O. mykiss* populations into a monophyletic lineage (including *O. mykiss nelsoni*) separate from most of the SMO populations and *O. clarkii*. However, some exceptions where observed. First, Río Acaponeta trout, the southernmost population, clusters with Mexican hatch-

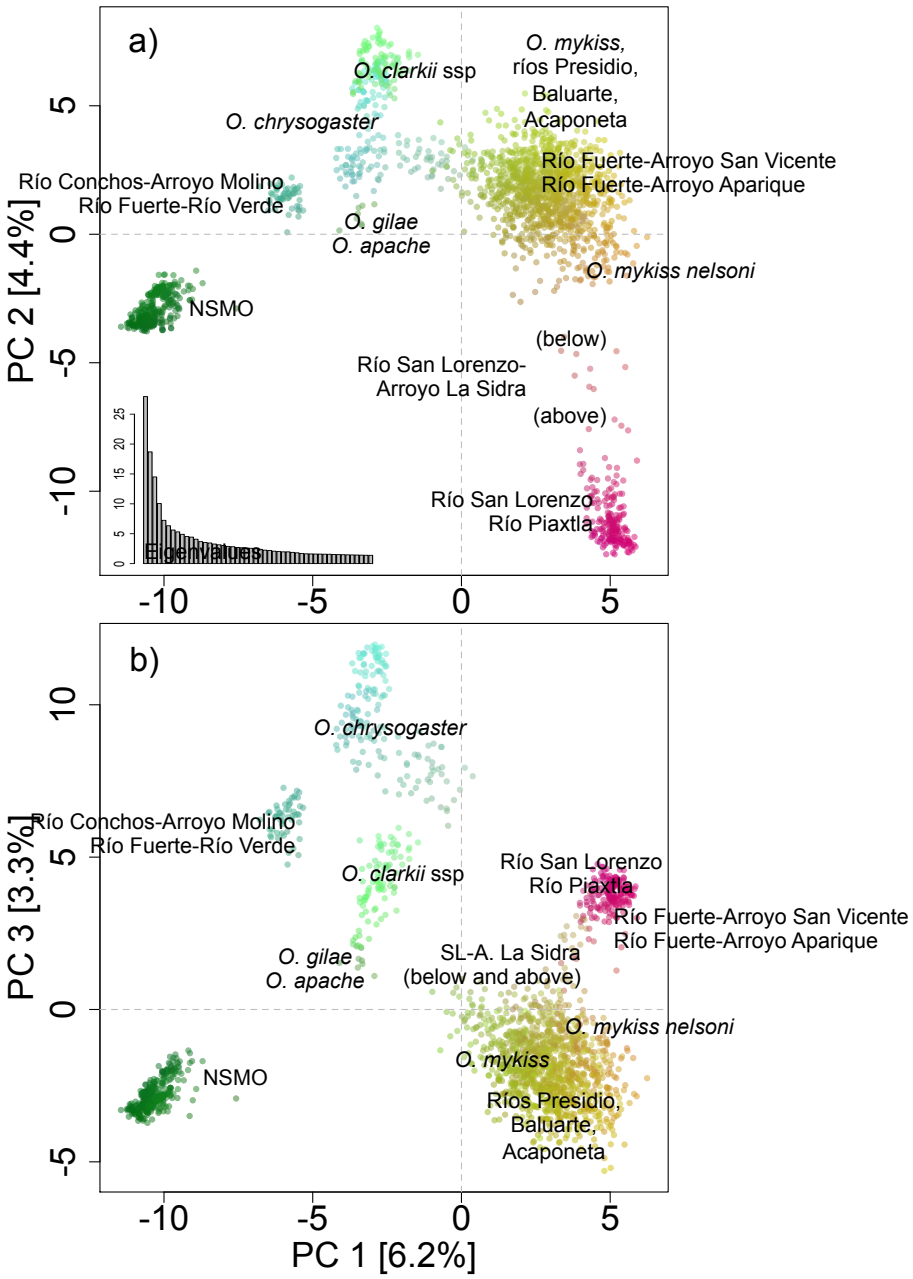

Figure 3.5: Principal Components Analysis (PCA) of allele frequencies from 18 microsatellites and 93 SNPs and the first 50 eigenvalues. a) First (PC1) and second (PC2) principal components and b) first (PC1) and third (PC3) principal components. The difference in color (red, blue and green channel) between clusters indicates divergence using the first three PCs. Seven clusters are shown (see text for description of cluster membership). NSMO: Northern Sierra Madre Occidental.

ery trout within the *O. mykiss* lineage in both trees and, second, two tributaries of Río Fuerte, the ríos Aparique and San Vicente, group with the Río San Lorenzo/Río Piaxtla lineage, although the branches grouping them do not have significant bootstrap support.

A long well-supported internal branch separates all the populations from ríos Yaqui, Mayo, and Guzmán as well as two tributaries of the Río Conchos. This result is consistent with the PCA. In addition, strong support was observed for a division between the northern and southern Río Yaqui regions. Populations from ríos Fuerte, Sinaloa and Culiacán (*O. chrysogaster*) form a monophyletic cluster on both trees, which also includes Río Conchos-Arroyo El Molino. Also consistent with the PCA, support was found for a cluster of ríos San Lorenzo and Piaxtla with populations interspersed with one another.

3.3.1 *Genetic introgression*

Results from the PCA and STRUCTURE analyses indicate that fish raised at all the Mexican hatcheries sampled in this study correspond to *O. mykiss* and are closely related to California hatchery rainbow trout strains (Figure 3.8 and Figure 3.9). I observed that introgression from hatchery rainbow trout is present in Mexican native trout populations. The analyses revealed that introgression is localized in tributaries where rainbow trout hatcheries occur, and that it varies from site to site (Figure 3.8 and Figure 3.9). Also, I observed completed shared ancestry between California hatchery rainbow trout and *O. chrysogaster* from the hatchery "Centro Trutícola Guachochi".

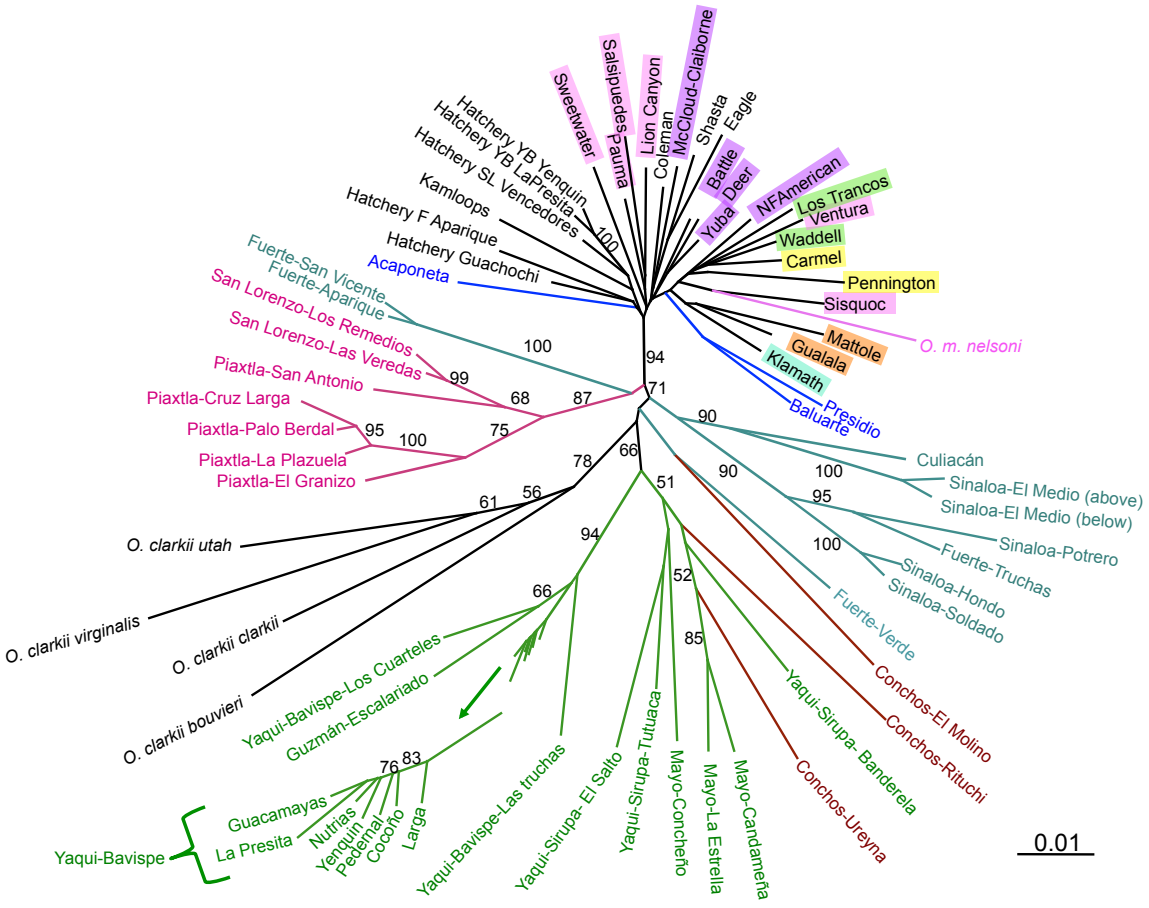

Figure 3.6: Unrooted neighbor-joining dendogram using data from 12 microsatellite markers. The tree was constructed with pairwise genetic distances and 1,000 bootstrapped distance matrices from 18 *O. mykiss*, 19 NSMO, 10 *O. chrysogaster*, and 12 SSMO natural-origin populations, four *O. clarkii* subspecies, and five Mexican and four California hatchery stocks. >50% percent bootstrap support is indicated in internal branches for Mexican trout (for *O. mykiss* see Garza *et al.* 2014). DPS affiliations of California *O. mykiss* populations (creeks) are highlighted in colors. Note that Battle Creek, Deer Creek and the Upper Yuba are not part of the Central Valley DPS but are part of the same region. Mexican natural-origin populations are indicated with branches and names colored.

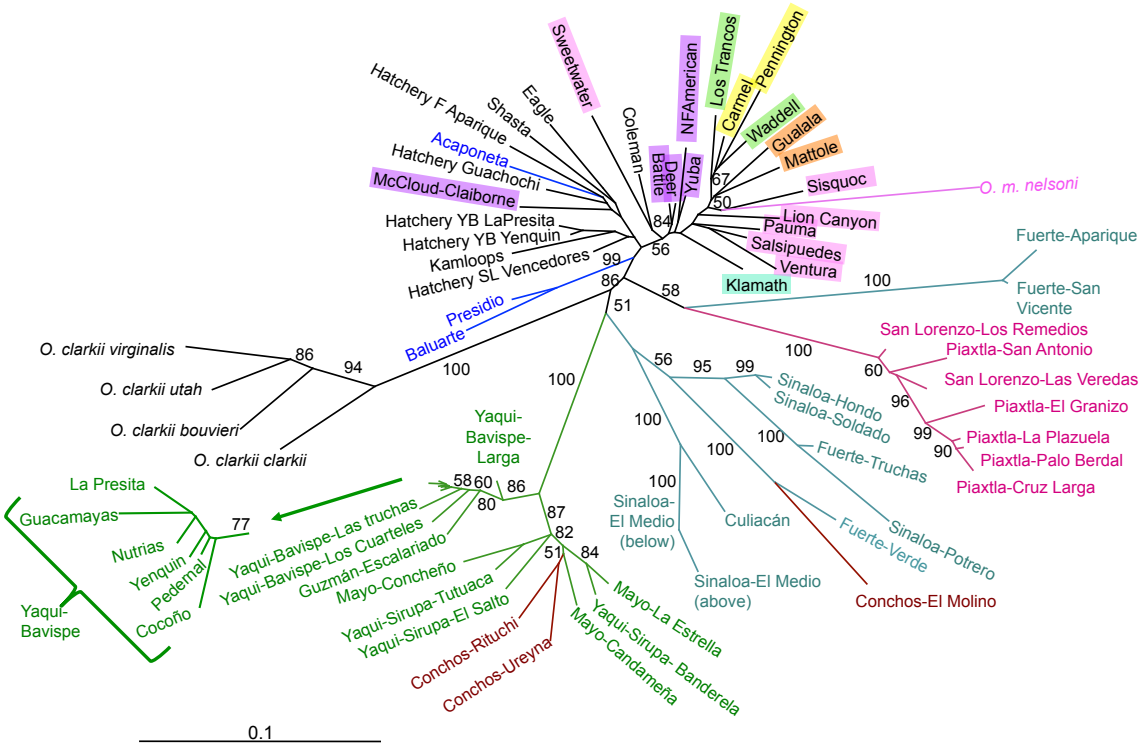

Figure 3.7: Unrooted neighbor-joining dendogram using both types of markers (12 microsatellites and 85 SNPs) combined. The dendogram was constructed with pairwise genetic distances and 1,000 bootstrapped distance matrices from the 18 *O. mykiss*, 19 NSMO, 10 *O. chrysogaster*, and 12 SSMO natural-origin populations, four *O. clarkii* subspecies, and five Mexican and four California hatchery stocks. >50% percent bootstrap support is indicated in internal branches. California *O. mykiss* populations (creeks) are highlighted in colors. Mexican natural-origin populations are indicated with branches and names colored.

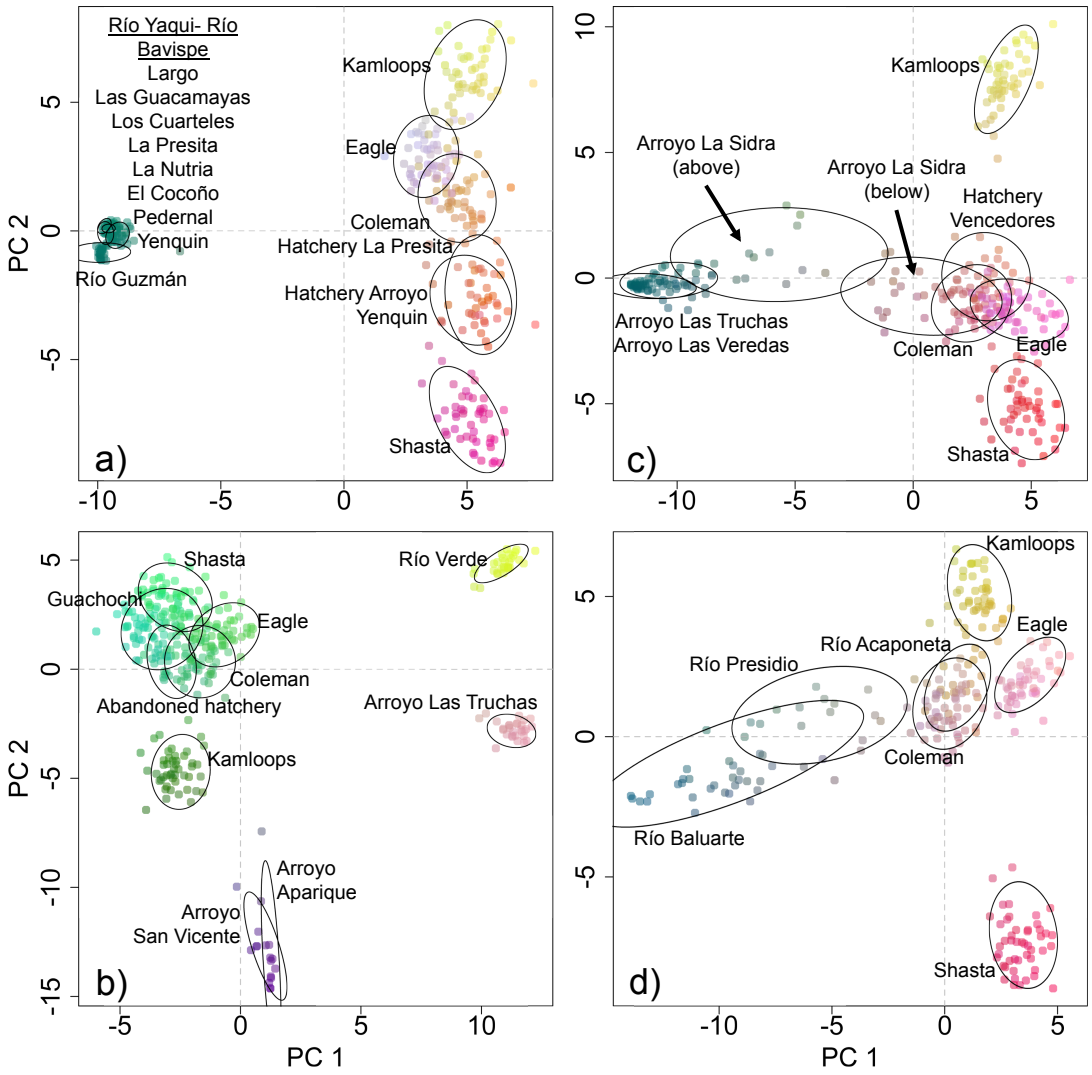

Figure 3.8: Principal Components Analysis (PCA) of allele frequencies. Each plot includes populations from four California hatchery strains (Coleman, Kamloops, Eagle Lake, and Mount Shasta) and a) Populations from all Río Yaqui-Río Bavispe and Río Guzmán tributaries, and samples from "Truchas La Presita" and "Yenquin" hatcheries; b) Populations from all Río Fuerte tributaries, samples from an abandoned hatchery located at the Río Fuerte-Arroyo Aparique, and samples from "Centro Trutícola Guachochi"; c) Populations from all Río San Lorenzo tributaries and samples from "Piscicultura Vencedores" hatchery; d) Populations from the three southernmost populations, ríos Presidio, Baluarte, and Acaponeta. The difference in color (red, blue and green channel) between clusters indicates divergence using the first three PCs.

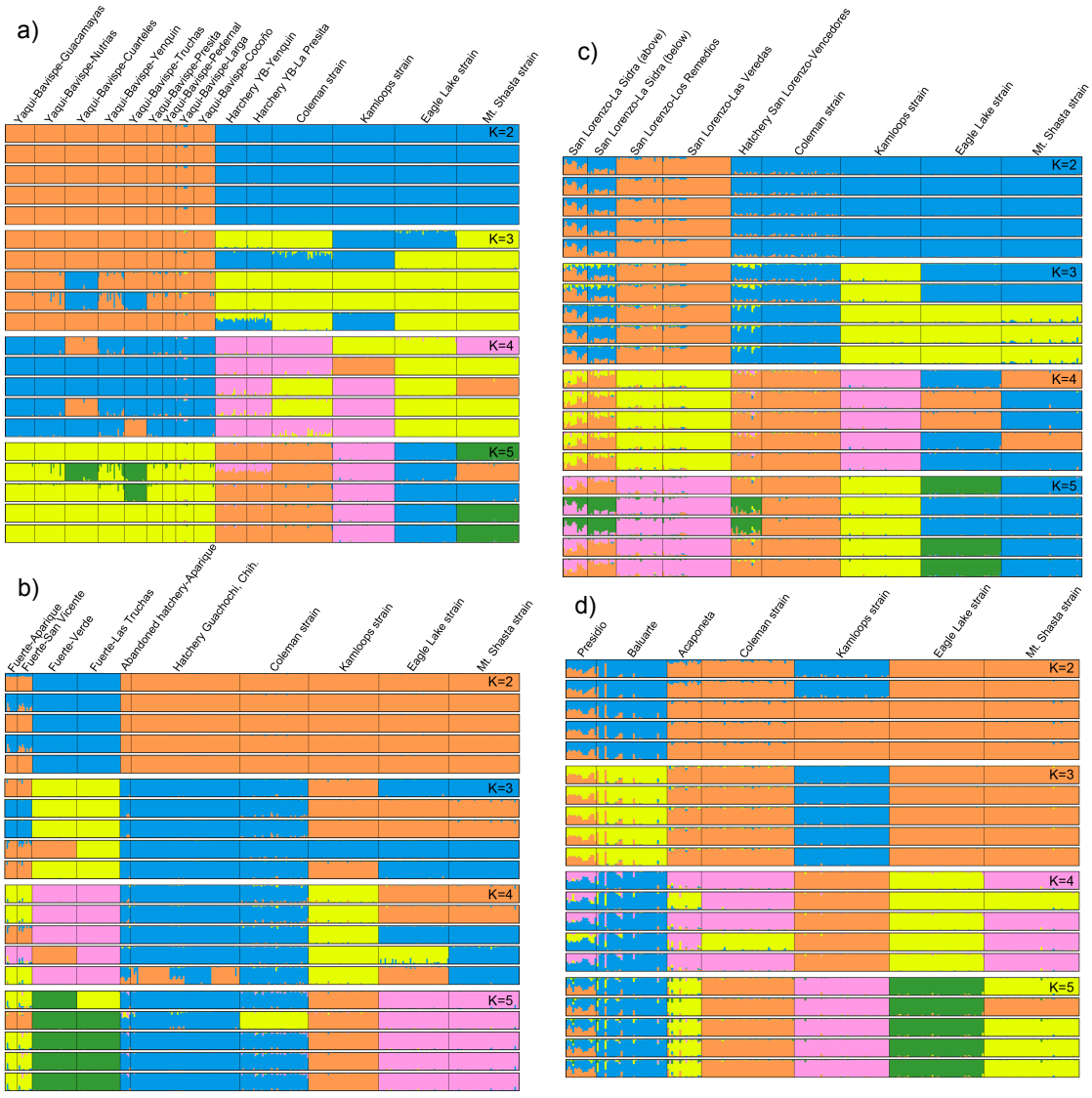

Figure 3.9: STRUCTURE analysis of hatchery rainbow trout ancestry. Estimated membership fraction (represented by color proportions) from four California hatchery strains (Coleman, Kamloops, Eagle Lake, and Mount Shasta) and a) Populations from all Río Yaqui-Río Bavispe and Río Guzmán tributaries, and samples from "Truchas La Presita" and "Yenquin" hatcheries; b) Populations from all Río Fuerte tributaries, samples from an abandoned hatchery located at the Río Fuerte-Arroyo Aparique, and samples from "Centro Trutícola Guachochi"; c) Populations from all Río San Lorenzo tributaries and samples from "Piscicultura Vencedores" hatchery; d) Populations from the three southernmost populations, ríos Presidio, Baluarte, and Acaponeta. Horizontal plots represent STRUCTURE runs constructed with Distruct. Each thin, colored, vertical line represents one individual. Populations are separated by vertical black lines. Five iterations of each K value are shown ($K = 2$ - 7).

3.4 Discussion

Despite previous efforts, little is known about the relationships among Mexican trout populations (Ruiz-Campos & Pister 1995; Hendrickson *et al.* 2002, 2006; Ruiz-Campos *et al.* 2003; Camarena-Rosales *et al.* 2007), and between them and other trout species (Nielsen *et al.* 1998; Nielsen & Sage 2001; De los Santos-Camarillo 2008; Mayden *et al.* 2010).

In this study, I find at least five major lineages of trout inhabiting northwestern Mexico that originated from at least two, and possibly three, separate colonization events. I found significant divergence between trout from the SMO and *O. mykiss* populations, as well as the other three previously named species analyzed (*O. apache*, *O. gilae*, and *O. clarkii*).

I confirmed the vast genetic diversity present in the trout inhabiting northwestern Mexico. In spite of the diversity previously shown in the Mexican trout complex, only two taxa have been formally described and protected by Mexican law (SEMARNAT 2000). These are Nelson's trout, *O. mykiss nelsoni* (Evermann 1908) and the Mexican golden trout, *O. chrysogaster* (Needham & Gard 1964). Nelson's trout is native to the Río Santo Domingo in the Sierra de San Pedro Mártir in northern Baja California (Evermann 1908; Snyder 1926; Ruiz-Campos & Pister 1995), but its taxonomic status as a subspecies has been questioned (Miller *et al.* 2005). Results from the STRUCTURE analysis and PCA indicate that Nelson's trout is, in fact, more closely related to *O. mykiss* than to SMO trout or other species, in agreement with early observations. This

subspecies was the first group of Mexican trout to be characterized, when Evermann (1908) described it as a new species (*Salmo nelsoni*). Later, Snyder (1926) concluded that these trout were closely related to *Salmo irideus*, now called *O. mykiss irideus* (Needham 1938). Moreover, the phylogenetic tree shows a stronger proximity of this group to *O. mykiss* Southern California DPS populations than to any other California populations or hatchery rainbow trout. This result indicates that during the most recent radiation of coastal steelhead, populations extended their range at least as far south as the Sierra de San Pedro Mártir in Baja California. Even though I did not find strong genetic differentiation of this population from other *O. mykiss* with the markers used, it is important to note that *O. m. nelsoni* had significantly lower genetic diversity than all other *O. mykiss* populations and the highest F_{ST} values among them (Tables S2 and S3), an indication of small effective population size (*Ne*) and long isolation time.

My results show extremely strong genetic differentiation among Mexican trout from the SMO, not only between basins but also at a smaller scale among localities within basins. I observed higher pairwise F_{ST} estimates between SMO trout (mean F_{ST} =0.351 - 0.684) than between *O. mykiss* populations (mean F_{ST} = 0.244 - 0.527). As mentioned above, F_{ST} is highly influenced by *Ne*. Small populations experience stronger effects of genetic drift and, in turn, reduced heterozygosity. This is directly related to the estimation of F_{ST}; therefore, the high values observed between SMO populations are likely a consequence of the small population sizes. This is also supported by the low microsatellite allelic richness and the proportion of polymorphic SNPs found in SMO trout.

Within the SMO trout, I found at least four well-differentiated lineages. The ríos Yaqui, Mayo, Guzmán, and the northern Río Conchos tributaries form a unique evolutionary unit, very different from the other species, as well as from other SMO drainages. This result is concordant with previous reports using mtDNA (Nielsen *et al.* 1998; Camarena-Rosales *et al.* 2007) and microsatellites (Nielsen & Sage 2001; De los Santos-Camarillo 2008). My analysis also confirmed the previously observed local structure within the Río Yaqui (Hendrickson *et al.* 1980; Nielsen & Sage 2001; Camarena-Rosales *et al.* 2007; De los Santos-Camarillo 2008), represented by the two main tributaries: the Río Bavispe populations to the north and the Río Sirupa ones to the south. Río Guzmán trout seem to be closely related to trout from Río Bavispe, while Río Mayo trout associate with Río Sirupa populations. These Río Bavispe/Río Guzmán and Río Sirupa/Río Mayo relationships were detected by Nielsen and Sage (2001), who discussed the hypotheses of multiple natural environmental events that interconnected several tributaries of the ríos Yaqui, Guzmán, Mayo, and Conchos basins (Hendrickson *et al.* 1980), permitting the migration of multiple species of fish from one to the other (Schönhuth *et al.* 2011; Domínguez-Domínguez *et al.* 2011), versus inter-basin transplants by humans (Behnke 1992). They concluded that both hypotheses are highly plausible and neither their nor my results can confirm one or the other and it could be a combination of both factors played a role in creating the observed patterns.

The results confirm that *O. chrysogaster* populations from ríos Fuerte, Sinaloa and Culiacán form a monophyletic group, with the exception of trout from two tributaries of the Río Fuerte (arroyos San Vicente and Aparique), which jumped from cluster

to cluster depending on the analysis used. For example, they cluster with *O. mykiss* or San Lorenzo/Piaxtla with the PCA depending on which principal components are used (Figure 3.5), but they cluster with San Lorenzo/Piaxtla on the phylogenetic trees (Figure 3.6 and Figure 3.7). This could be the result of reduced genetic variation or genetic introgression. Also, I found very strong genetic differentiation between and within the three basins where they were known to occur. However, I did not observe any evidence of subgrouping by basin as in Ruiz-Campos *et al.* (2003) and Camarena-Rosales *et al.* (2007), but a strong association between tributaries from different basins that are geographically adjacent to each other (PCA not shown), an indication of migration between them. More detailed studies would be necessary to understand the small-scale structure found here.

Río Conchos trout were originally described as "cutthroat type" (Cope 1886) but then not seen for decades. Trout were recently rediscovered in the Río Conchos after exhaustive efforts from the group "Truchas Mexicanas" (Hendrickson *et al.* 2006). My results do not indicate that populations from this basin are related to cutthroat trout, but more closely related to either the Yaqui/Mayo/Guzmán complex or to *O. chrysogaster*. Trout from ríos Rituchi and Ureyna, tributaries of the Río Conchos, group tightly with Río Yaqui-Río Bavispe trout. In contrast, Conchos-El Molino and Fuerte-Verde formed a single cluster on the PCA and cluster together in the trees along with the *O. chrysogaster* lineage. In spite of the fact that these two tributaries are on alternative sides of the continental divide, they are geographically adjacent (Figure 3.1) and stream capture episodes could have caused dispersal from Río Fuerte-Verde into

Río Conchos-El Molino. This movement of fish between Río Fuerte and Río Conchos has been previously reported in other freshwater fish species (Schönhuth *et al.* 2011, 2014; Domínguez-Domínguez *et al.* 2011).

Río San Lorenzo and Río Piaxtla formed another independent evolutionary group in the analyses. Unfortunately, trout from Río San Lorenzo-Arroyo La Sidra are heavily introgressed by hatchery rainbow trout (Figure 3.8 and Figure 3.9), as indicated in previous reports where a considerable number of migrants between native and exotic trout ($Nm = 2.7$) was estimated (De los Santos-Camarillo 2008). Samples taken below the hatchery (downstream) showed a higher level of introgression than those above the hatchery (upstream). No introgression was detected at other Río San Lorenzo tributaries, indicating that hatchery rainbow trout has not extended its range in this basin beyond the immediate vicinity of the hatchery.

I observed that trout from ríos Presidio, Baluarte and Acaponeta are more closely related to *O. mykiss* than the other SMO lineages. Hybridization between trout from these southern drainages and exotic rainbow trout has been previously detected (De los Santos-Camarillo 2008) and other studies have considered them to be introduced rainbow trout based on morphologic characteristics (Miller *et al.* 2005). My results indicate that trout from ríos Presidio and Baluarte may be partly of hatchery origin, although the combined tree (Figure 3.7), the PCA (Figure 3.8) and the STRUCTURE runs (Figure 3.9) suggest that they may form a unique lineage of, at least partially, native ancestry. These observations could be the result of a more recent natural colonization event of an anadromous *O. mykiss* or hybridization between native trout and

hatchery rainbow trout. Previous studies that looked at impacts of hatchery rainbow trout on natural origin trout in southern California indicate that they did not have a significant impact on the naturally spawning populations (Clemento *et al.* 2009) and suggested that this low contribution of hatchery fish to natural reproduction could be the consequence of ancestral differences in reproductive patterns. Natural origin trout from the Presidio, Baluarte and Acaponeta are at the extreme south of the species natural distribution and they are presumably well adapted to the local conditions prevalent in their environment. This local adaptation could be somewhat acting as a reproductive barrier between the native and exotic trout, explaining the divergence observed between trout from the ríos Presidio and Baluarte. Also, strong differences in body size between native and introduced rainbow trout in that region has been suggested as an impediment to hybridization (De los Santos-Camarillo 2008). Unfortunately, trout from Río Acaponeta clusters with *O. mykiss* in every analysis performed here and, more specifically, this trout associates with hatchery rainbow trout strains. These results provide strong evidence that trout from Río Acaponeta are descended directly from hatchery rainbow trout or, if still present, that the native population is completely introgressed.

Samples from Centro Trutícola, the hatchery in Guachochi, Chihuahua come from a program with objectives to raise and cross a strain of *O. chrysogaster* for conservation purposes (Barriga-Sosa *et al.*). Unfortunately, my results indicate that trout from this hatchery do not correspond to *O. chrysogaster* or any other native SMO lineages. In contrast, they tightly cluster with hatchery rainbow trout, indicating incorrect identification of the fish when collected, or total introgression from rainbow trout also

raised at the hatchery.

The first documented introduction of non-native trout into Mexican waters was in 1886, when about 33,000 *O. mykiss irideus* eggs were imported from Baird Station on the McCloud River, California, United States (Arredondo-Figueroa 1983). The total number of rainbow trout hatcheries in Mexico is unknown, but some unofficial reports indicate that there are around 40 hatcheries that produce 110 tons a year in the state of Durango alone and about 182 hatcheries that produce 184 tons a year in the state of Chihuahua (Diaz 2010; Aquahoy 2011). Both of these states possess native trout populations.

Fish reared at hatcheries face problems such as domestication selection, inbreeding depression and increased disease susceptibility. Introductions of non-native species in any environment can have devastating effects on local species. These effects can range from reduction of the native genetic diversity, to complete extinction of local populations (Rhymer & Simberloff 1996; USFWS 2003). Introgression from non-native rainbow trout was present at different levels in most of the tributaries with established hatcheries, but the genetic integrity of native trout from northwestern Mexico is still maintained in many watersheds. The information in the present study is crucial to guide effective conservation strategies for this globally important group of fishes.

3.5 References

Abadía-Cardoso A, Clemento AJ, Garza JC (2011) Discovery and characterization of single nucleotide polymorphisms in steelhead/rainbow trout, *Oncorhynchus mykiss*. Molecular Ecology Resources 11, 31-49.

Aguilar A, Garza JC (2006) A comparison of variability and population structure for major histocompatibility complex and microsatellite loci in California coastal steelhead (*Oncorhynchus mykiss* Walbaum). Molecular Ecology 15, 923-937.

Aguilar A, Garza JC (2008) Isolation of 15 single nucleotide polymorphisms from coastal steelhead, *Oncorhynchus mykiss* (Salmonidae). Molecular Ecology Resources 8, 659-662.

Allendorf FW, Hitt NP, Knudsen KL, Lundquist LL, Spruell P (2004) Intercrosses and the U.S. Endangered Species Act: Should hybridized populations be included as West-slope cutthroat trout? 18, 1203-1213.

Aquahoy (2011) México: Gobierno de Durango impulsa la producción y el consumo de la trucha arco-iris.

Arredondo-Figueroa J (1983) Especies animales acuáticas de importancia nutricional, introducidas en México. Biotica, 9, 2339.

Banks MA, Blouin MS, Baldwin BA, *et al.* (1999) Isolation and inheritance of novel microsatellites in Chinook salmon (*Oncorhynchus tschawytscha*). Journal of Heredity 90, 281-288.

Barriga-Sosa IA, Arredondo-Figueroa JL, De-La-Mora GI, Banda-Cortés M, *et al.* (n.d.) Primeras experiencias del cultivo de la trucha dorada (*Oncorhynchus chrysogaster*) Retrieved from http://www.izt.uam.mx/pexpa/htms/proyect3.htm

Behnke RL (1992) *Native trout of western North America.* American Fisheries Society Monograph 6, Bethesda.

Behnke RJ (2002) *Trout and salmon of North America.* Chanticleer Press, Inc., New York.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Universit de Montpellier II, Montpellier, France.

Bjorkstedt EP, Spence BC, Garza JC, *et al.* (2005) An analysis of historical population

structure for evolutionarily significant units of Chinook salmon, coho salmon, and steelhead in the north-central California coast recovery domain, p. 210. NOAA Technical Memorandum NMFS (NOAA-TM-NMFS-SWFSC-382).

Busby PJ, Wainwright TC, Bryant GJ, *et al.* (1996) Status review of West coast steelhead from Washington, Idaho, Oregon, and California. NOAA Technical Memorandum NMFS-NWFSC-27. US Department of Commerce.

Camarena-Rosales F, Ruiz-Campos G, De La Rosa-Velez J, *et al.* (2007) Mitochondrial haplotype variation in wild trout populations (Teleostei : Salmonidae) from northwestern Mexico. Reviews in Fish Biology and Fisheries 18, 33-45.

Campbell NR, Overturf K, Narum SR (2009) Characterization of 22 novel single nucleotide polymorphism markers in steelhead and rainbow trout. Molecular Ecology Resources 9, 318-322.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis models and estimation procedures. Evolution 21, 550-570.

Cavender TM, Miller RR (1982) *Salmo australis*, a new species of fossil sal-monid from southwestern Mexico. In: Contributions from the museum of paleontology (ed. Smith GR), pp. 1-17. The University of Michigan.

Clemento AJ, Anderson EC, Boughton D, Girman D, Garza JC (2009) Population genetic structure and ancestry of *Oncorhynchus mykiss* populations above and below dams in south-central California. Conservation Genetics 10, 1321-1336.

Condrey MJ, Bentzen P (1998) Characterization of coastal cutthroat trout (*Oncorhynchus clarki clarki*) microsatellites and their conservation in other salmonids. Molecular Ecology 7, 787-789.

Cope ED (1886) The most southern salmon. American Naturalist, 20, 735.

Diaz L (2010) Compra Chihuahua el 70% del pescado que consume. El Diario de Chihuahua.

Domínguez-Domínguez O, Vila M, Pérez-Rodríguez R, Remón N, Doadrio I (2011) Complex evolutionary history of the Mexican stoneroller *Campostoma ornatum* Girard, 1856 (Actinopterygii: Cyprinidae). BMC Evolutionary Biology, 11, 153.

De Los Santos-Camarillo AB (2008) Definición de unidades taxonómicas en el complejo de truchas del noroeste de México, mediante el análisis de marcadores microsatálites, Masters thesis. Centro de Investigaciones Biológicas del Noreste.

Dowling TE, Childs MR (1992) Impact of hybridization on a threatened trout of the southwestern United States. Conservation Biology, 6, 355-364.

Evermann BW (1908) Descriptions of a new species of trout (*Salmo nelsoni*) and a new Cyprinodont (*Fundulus meeki*) with notes on other fishes from lower California. Proceedings of the Biological Society of Washington 21, 19-30.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10, 564-567.

Felsenstein J (2005) PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Garza JC, Gilbert-Horvath EA, Spence BC, *et al.* (2014) Population structure of steelhead in coastal California. Transactions of the American Fisheries Society 143, 134-152.

Garza JC, Pearse DE (2008) Population genetic structure of *Oncorhynchus mykiss* in the California Central Valley. Final report for California Department of Fish and Game Contract No. PO485303, p. 52. NOAA, Southwest Fisheries Science Center.

Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical F-statistics. Molecular Ecology Notes 5, 184-186.

Hedrick PW (1999) Perspective: Highly variable loci and their interpretation in evolution and conservation. Evolution 53, 313-318.

Hendrickson DA, Espinosa-Pérez H, Findley LT, *et al.* (2002) Mexican native trouts: a review of their history and current systematic and conservation status. Reviews in Fish Biology and Fisheries 12, 273-316.

Hendrickson DA, Minckley WL, Miller RR, Siebart DJ, Minckley PH (1980) Fishes of the Río Yaqui Basin, México and United States. Journal of the Arizona-Nevada Academy of Science 15, 65-106.

Hendrickson DA, Neely DA, Mayden RL, *et al.* (2006) Conservation of Mexican native trout and the discovery, status, protection and recovery of the Conchos trout, the first native *Oncorhynchus* of the Atlantic drainage in Mexico. In: Studies of North American desert fishes in honor of E. P. (Phil) Pister, conservationist. (eds. Lozano-Vilano MdL, Contreras-Balderas AJ), pp. 162-201. Universidad Autonoma de Nuevo León, Mexico.

IUCN (2010) IUCN red list of threatened species. Version 2010.4. Downloaded on 21

February 2011 from iucnredlist.org.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23, 18011806.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403-1405.

Leary R, Allendorf FW, Knudsen KL (1985) Developmental instability and high meristic counts in interspecific hybrids of salmonid fishes. Evolution, 39, 13181326.

Loudenslager EJ, Rinne JN, Gall GAE, David RE (1986) Biochemical genetic studies of native Arizona and New Mexico trout. Southwestern Naturalist 31, 221-234.

Mayden RL, Dillman CB, Espinosa-Pérez H, *et al.* (2010) Evolution and diversity of trout species in the Sierra Madre Occidental of Mexico. In: Conserving Wild Trout. Proceedings of the Wild Trout X symposium (eds Carline RF, LoSapio C), pp. 134144. Bozeman, Montana.

Mayden RL, Wood RM (1995) Systematics, species concepts, and the evolutionarily significant unit in biodiversity and conservation biology. In: Evolution and the aquatic ecosystem: defining unique units in population conservation. Symposium 17 (ed. Nielsen JL), pp. 58113. American Fisheries Society, Bethesda, Maryland.

McConnell S, Hamilton L, Morris DB, *et al.* (1995a) Isolation of salmonid microsatellite loci and their application to the population genetics of Canadian east coast stocks of Atlantic salmon. Aquaculture 137, 19-30.

McConnell S, O'Reilly P, Hamilton L, Wright JM, Bentzen P (1995b) Polymorphic microsatellite loci from Atlantic salmon (*Salmo salar*): genetic differentiation of North American and European populations. Canadian Journal of Fisheries and Aquatic Sciences 52, 1863-1872

Miller RR, Minckley WL, Norris SM (2005) Freshwater fishes of Mexico. The University of Chicago Press, Chicago and London.

Moraes CT, Andreetta F, Bonilla E *et al.* (1991) Replication-competent human mitochondrial DNA lacking the heavy-strand promoter region. Molecular and Cellular Biology 11, 16311637.

Morris DB, Richard KR, Wright JM (1996) Microsatellites from rainbow trout (*Oncorhynchus mykiss*) and their use for genetic study of salmonids. Canadian Journal of

Fisheries and Aquatic Sciences 53, 120-126.

Needham PR (1938) Notes on the introduction of *Salmo nelsoni* (Evermann) into California from Mexico. Transactions of the American Fisheries Society 67, 139-146.

Needham PR, Gard R (1964) A new trout from central Mexico: *Salmo chrysogaster*, the Mexican golden trout. Copeia 1964, 169-173.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89, 583-590.

Nielsen JL, Fountain MC, Favela JC, Cobble K, Jensen BL (1998) *Oncorhynchus* at the southern extent of their range: a study of mtDNA control-region sequence with special reference to an undescribed subspecies of *O. mykiss* from Mexico. Environmental Biology of Fishes 51, 7-23.

Nielsen JL, Fountain MC, Wright JM (1997) Biogeographic analysis of Pacific trout (*Oncorhynchus mykiss*) in California and Mexico based on mtDNA and nuclear microsatellites. In: Molecular Systematics of Fishes (eds. Kocher TD, Stepien CA), pp. 53-73. Academic Press, San Diego.

Nielsen JL, Sage GK (2001) Microsatellite analyses of the trout of northwest Mexico. Genetica 111, 269-278.

Nielsen JL, Sage GK (2002) Population genetic structure in Lahontan cutthroat trout. Transactions of the American Fisheries Society 131, 376-388.

O'Reilly PT, Hamilton LC, McConnell SK, Wright JM (1996) Rapid analysis of genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. Canadian Journal of Fisheries and Aquatic Sciences 53, 2292-2298.

Pearse DE, Donohoe CJ, Garza JC (2007) Population genetics of steelhead (*Oncorhynchus mykiss*) in the Klamath River. Environmental Biology of Fishes 80, 377-387.

Phillips R, Ráb P (2001) Chromosome evolution in the Salmonidae (Pisces): an update. Biological Reviews 76, 1-25.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155, 945-959.

R Development Core Team (2011) A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. Annual Review of Ecology and Systematics, 27, 83109.

Rosenberg N (2004) Distruct: a program for the graphical display of population structure. Molecular Ecology Notes 4, 137138.

Rousset F (2008) GENEPOP' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources, 8, 103106.

Rubidge E, Corbett P, Taylor EB (2001) A molecular analysis of hybridization between native westslope cutthroat trout and introduced rainbow trout in southeastern British Columbia, Canada. Journal of Fish Biology, 59, 4254.

Ruiz-Campos G, Camarena-Rosales F, Varela-Romero A, Sánchez-Gonzáles S, De La Rosa-Vélez J (2003) Morphometric variation of wild trout populations from northwestern Mexico (Pisces : Salmonidae). Reviews in Fish Biology and Fisheries 13, 91-110.

Ruiz-Campos G, Pister EP (1995) Distribution, habitat, and current status of the San Pedro Martir Rainbow Trout, Oncorhynchus mykiss nelsoni (Evermann). Bulletin Southern California Academy of Sciences 94, 131-148.

Schönhuth S, Blum MJ, Lozano-Vilano L et al. (2011) Inter-basin exchange and repeated headwater capture across the Sierra Madre Occidental inferred from the phylogeography of Mexican stonerollers. Journal of Biogeography, 38, 1406-1421.

Schönhuth S, Perdices A, Lozano-Vilano L et al. (2014) Phylogenetic relationships of North American western chubs of the genus *Gila* (Cyprinidae, Teleostei), with emphasis on southern species. Molecular Phylogenetics and Evolution 70, 210-230.

Scribner KT, Gust JR, Fields RL (1996) Isolation and characterization of novel salmon microsatellite loci: Cross-species amplification and population genetic applications. Canadian Journal of Fisheries and Aquatic Sciences 53, 833-841.

SEMARNAT (2000) Norma oficial mexicana NOM- 059-ECOL- 2001, protección ambiental-especies nativas de México de flora y fauna silvestres- Categorías de riesgo y especificaciones para su inclusión, exclusión o cambio- Lista de especies en riego, SEMARNAT (Secretaria del Medio Ambiente y Recursos Naturales) Diario Oficial, 6 de Marzo.

Shepard BB, May BE, Urie W (2005) Status and conservation of westslope cutthroat trout within the western United States. North American Journal of Fisheries Management, 25, 1426-1440.

Small MP, Beacham TD, Withler RE, Nelson RJ (1998) Discriminating coho salmon (*Oncorhynchus kisutch*) populations within the Fraser River, British Columbia, using microsatellite DNA markers. Molecular Ecology 7, 141-155.

Smith CT, Koop BF, Nelson RJ (1998) Isolation and characterization of Coho salmon (*Oncorhynchus kisutch*) microsatellites and their use in other salmonids. Molecular Ecology 7, 1614-1616.

Snyder JO (1926) The trout of the Sierra San Pedro Mártir, Lower California. University of California Publ Zool 21, 419-426.

USFWS (2003) *Gila trout recovery plan.* 3rd revision. Albuquerque.

Weigel DE, Peterson JT, Spruell P (2003) Introgressive hybridization between native cutthroat trout and introduced rainbow trout. Ecological Applications, 13, 3850.

Wenburg JK, Bentzen P, Foote CJ (1998) Microsatellite analysis of genetic population structure in an endangered salmonid: the coastal cutthroat trout (*Oncorhynchus clarki clarki*). Molecular Ecology 7, 733-749.

Williamson KS, Cordes JF, May B (2002) Characterization of microsatellite loci in Chinook salmon (*Oncorhynchus tshawytscha*) and cross-species amplification in other salmonids. Molecular Ecology Notes 2, 17-19.

Wilson WD, Turner TF (2009) Phylogenetic analysis of the Pacific cutthroat trout (*Oncorhynchus clarki* ssp.: Salmonidae) based on partial mtDNA ND4 sequences: A closer look at the highly fragmented inland species. Molecular Phylogenetics and Evolution, 52, 406415.

General Conclusion

The dissertation presented here represents an in-depth evaluation of trout populations in Northwest America at different scales, from the evaluation of biological traits throughout the reconstruction of pedigrees in two populations to a phylogeographic examination of multiple trout species. Here, the combination of novel molecular techniques allow me to answer critical ecological questions for the appropriate management of perhaps one of the most important group of fish in the world: The Pacific Trout.

Chapter one is a description of the discovery, characterization and development of assays for a large number (139) of SNP loci for steelhead/rainbow trout (Abadía-Cardoso *et al.* 2011). I exploited EST databases to design nearly 500 primer sets for functional genome regions. PCR products resulting from these genes, which include both exonic and intronic regions, were then sequenced in an ascertainment panel of 22 fish designed to simultaneously represent some of the phylogenetic diversity of the species and to provide polymorphic markers for focal populations in California. These SNP markers represent a valuable resource for studying ecological interactions, phylogeography, and conservation status, as well as for pedigree reconstruction, individual and genetic stock identification and, eventually, for linkage mapping.

During the past decade, microsatellite markers have dominated population genetic work in salmonids, due to their high variability and conservation among related species (Aguilar & Garza 2006; Clemento *et al.* 2009; Pearse *et al.* 2007; Pearse *et al.* 2009). However, microsatellites have significant drawbacks, among them relatively high genotyping error/mutation rates, significant staff time necessary for data generation and allele calling, and homoplasy. Moreover, the results obtained with microsatellites

in one laboratory are not directly combinable with data generated in other laboratories (Seeb *et al.* 2007). Conversely, data obtained from SNP loci are easily portable and combinable between laboratories. Although SNPs were initially granted dim prospects for relationship inference in molecular ecology (Glaubitz *et al.* 2003), Anderson and Garza (2006) demonstrated that a relatively small number of SNPs (< 100) would allow accurate parentage studies larger than any that had been previously attempted. The coincidence of that work with the advent of novel genotyping platforms that permit the rapid genotyping of thousands of individuals at many loci has now set the stage for SNPs to be the marker of choice for large-scale parentage studies and for genetic tagging of migratory species.

Chapter two consists of the elucidation of critical reproductive patterns in ESA-listed steelhead from a supplementation program in the Russian River, CA, using a pedigree-based intergenerational genetic tagging protocol to provide information comparable to that obtained by physical tagging methods (Abadía-Cardoso *et al.* 2013). Artificial propagation and subsequent supplementation can have numerous negative effects on natural populations (Utter 1998; Bryant & Reed 1999; Frankham 2008; Williams & Hoffman 2009; Christie *et al.* 2012) and detailed estimates of reproductive and behavioral trait values of the propagated population is a critical first step in understanding and minimizing these consequences. I demonstrated how the use of pedigree-based genetic tagging provides a powerful means of understanding many basic biological traits in relatively high fecundity species with significant conservation concerns. The use of such analyses as a surrogate for traditional tagging methods provided us with a large

number of pedigrees, and allowed us to evaluate patterns on a family level. Specifically, I assigned most individuals that return from the ocean to pairs of parents that were spawned on the same day, but without cross information recorded. A matching samples analysis allowed us to estimate the number of fish that were spawned multiple times within a single season and the number that return and reproduce in multiple seasons. These estimates reveled inconsistencies from hatchery program goals. The two hatchery programs evaluated here use local fish as broodstock, provide substantial numbers of spawners in natural areas, and are not genetically differentiated from the natural populations in the Russian River (Deiner *et al.* 2007). As such, elucidation of life history patterns in these hatchery steelhead populations allowed us to examine whether they were negatively influencing the associated natural populations. I estimated the variation in family size and the age distribution amongst reproducing fish. I observed a high proportion of age-two spawners, contrasting with the management plan for these hatchery programs (FISHPRO 2004) that recommends less than 1% of spawners be age-two fish. I also found that fish of different ages spawn on significantly different dates. These patterns on a family level are exactly what is necessary for a classical estimation of trait heritability using parent/offspring and sibling/sibling regression (Fisher 1918). The finding that spawning time in steelhead is highly heritable is both a novel and important outcome of this approach, with implications for management and conservation. All the information obtained in this chapter will allow for better strategies for supplementation programs and ultimately lead to more effective conservation and management plans.

Simultaneously, the powerful molecular tools developed along with the extensive sampling effort, were applied in chapter three to population genetic analysis of the Mexican trout, to evaluate population structure and differentiation, and to understand its phylogeographic distribution. An important first steps in construction of an effective conservation strategy for any taxon is to document the diversity of biological units in that taxon and gain understanding of the evolutionary processes that result in the generation of those biological units (Mayden & Wood 1995). However, the taxonomic status of native trout inhabiting northwestern Mexico has been the subject of speculation and controversy for decades. Only two taxa from the Mexican trout complex have been formally described, even though the complex has long been considerd as highly diverse (Behnke 2002), and other groups of trout distributed along the Sierra Madre Occidental (SMO) have not been formally described. For years, researchers have tried to investigate the evolutionary relationships among the Mexican trout complex providing valuable information and some insight into the diversity of the trout inhabiting northwestern Mexico. However, these groups of trout have been considered as undescribed subspecies of *O. mykiss* (Nielsen & Sage 2001; Behnke 2002). It is evident that the incomplete sampling effort, the small sample sizes, and the low resolution of these analyses have left many unresolved questions. Conservation of the trout inhabiting northwestern Mexico first requires complete documentation of the genetic diversity, as well as a complete understanding of the evolutionary history of this species complex. My analysis included fish from all the Mexican basins in which native trout have been reported. I also put these Mexican populations in a phylogeographic perspective by

comparing them with other trout species (*O. mykiss* and *O. clarkii*), specifically with those that are presumably closely related and with a similar evolutionary history (*O. gilae* and *O. apache*).

Here, I confirmed the vast genetic diversity present in the trout inhabiting northwestern Mexico. I also confirmed that Nelson's trout is, in fact, more closely related to *O. mykiss* than to SMO trout or other species. I provided evidence that trout inhabiting the SMO correspond to independent lineages separated from *O. mykiss*. The results show extremely strong genetic differentiation among Mexican trout from the SMO, not only between basins but also at a smaller scale among localities within basins. Within the SMO trout I found at least four well-differentiated lineages. The ríos Yaqui, Mayo, and Guzmán form a unique evolutionary unit and the ríos San Lorenzo and Piaxtla another one, very different from the other species. *O. chrysogaster* populations form a monophyletic group and show strong differentiation between and within the three basins where they were known to occur. Río Conchos trout were originally described as "cutthroat type" but then not seen for decades. Trout were recently rediscovered in the Río Conchos after exhaustive efforts from the group "Truchas Mexicanas" (Hendrickson *et al.* 2006). The present work represents the first to genetically examine Río Conchos trout since these small populations were rediscovered. My results do not indicate that populations from this basin are related to cutthroat trout, but alternatively that populations from tributaries of the northern Río Conchos are more closely related to the Yaqui/Mayo/Guzmán complex and tributaries from the southern Río Conchos to *O. chrysogaster*. Finally, I observed that trout from ríos Presidio, Baluarte and Acaponeta

are more closely related to *O. mykiss* than the other SMO lineages, and it is unclear if it is a result of hybridization with hatchery rainbow trout or a more recent natural colonization event of an anadromous *O. mykiss*.

Introgression from non-native rainbow trout was present at different levels in most of the tributaries with established hatcheries, but the genetic integrity of native trout from Northwestern Mexico is still maintained in many watersheds. The information in the present study is crucial to guide effective conservation strategies for this globally important group of fishes.

References

Abadía-Cardoso A, Anderson EC, Pearse DE, Garza JC (2013) Large-scale parentage analysis reveals reproductive patterns and heritability of spawn timing in a hatchery population of steelhead (*Oncorhynchus mykiss*). Molecular Ecology 22, 4733-4746.

Abadía-Cardoso A, Clemento AJ, Garza JC (2011) Discovery and characterization of single nucleotide polymorphisms in steelhead/rainbow trout, *Oncorhynchus mykiss*. Molecular Ecology Resources 11, 31-49.

Aguilar A, Garza JC (2006) A comparison of variability and population structure for major histocompatibility complex and microsatellite loci in California coastal steelhead (*Oncorhynchus mykiss* Walbaum). Molecular Ecology 15, 923-937.

Behnke RJ (2002) Trout and salmon of North America Chanticleer press, inc., New York.

Bryant EH, Reed DH (1999) Fitness decline under relaxed selection in captive populations. Conservation Biology 13, 665-669.

Christie MR, Marine ML, French RA, Blouin MS (2012) Genetic adaptation to captivity can occur in a single generation. Proceedings of the National Academy of Sciences of the United States of America 109, 238-242.

Clemento AJ, Anderson EC, Boughton D, Girman D, Garza JC (2009) Population genetic structure and ancestry of *Oncorhynchus mykiss* populations above and below dams

in south-central California. Conservation Genetics 10, 1321-1336.

Deiner K, Garza JC, Coey R, Girman DJ (2007) Population structure and genetic diversity of trout (*Oncorhynchus mykiss*) above and below natural and man-made barriers in the Russian River, California. Conservation Genetics 8, 437-454.

Fisher RA (1918) The correlation between relatives on the supposition of Mendelian inheritance. Philosophical Transactions of the Royal Society of Edinburgh 52, 399433.

FISHPRO (2004) Hatchery and genetic management plans for Russian River fish production facilities Coho salmon and steelhead., p. 73. FISHPRO, a Division of HDR and ENTRIX, inc. Available from: nrm.dfg.ca.gov.

Frankham R (2008) Genetic adaptation to captivity in species conservation programs. Molecular Ecology 17, 325-333.

Glaubitz JC, Rhodes OE, Dewoody JA (2003) Prospects for inferring pairwise relationships with single nucleotide polymorphisms. Molecular Ecology 12, 1039-1047.

Mayden RL, Wood RM (1995) Systematics, species concepts, and the evolutionarily significant unit in biodiversity and conservation biology. In: Evolution and the aquatic ecosystem: defining unique units in population conservation. Symposium 17 (ed. Nielsen JL), pp. 58-113. American Fisheries Society, Bethesda, Maryland.

Nielsen JL, Sage GK (2001) Microsatellite analyses of the trout of northwest Mexico. Genetica 111, 269-278.

Pearse DE, Donohoe CJ, Garza JC (2007) Population genetics of steelhead (*Oncorhynchus mykiss*) in the Klamath River. Environmental Biology of Fishes 80, 377-387.

Pearse DE, Hayes SA, Bond MH, *et al.* (2009) Over the falls? Rapid evolution of ecotypic differentiation in steelhead/rainbow trout (*Oncorhynchus mykiss*). Journal of Heredity 100, 515-525.

Seeb LW, Antonovich A, Banks MA, *et al.* (2007) Development of a standardized DNA database for Chinook salmon. Fisheries 32, 540-552.

Utter FM, Allendor FW, Hodgins HO (1973) Genetic variability and relationships in Pacific salmon and related trout based on protein variations. Systematic Zoology 22, 257-270.

Williams SE, Hoffman EA (2009) Minimizing genetic adaptation in captive breeding programs: a review. Biological Conservation 142, 2388-2400.

Supplemental material

Supplement 3.1. Pairwise F_{ST} estimates for all populations and strains (below diagonal) and significance p-values (above diagonal). Labels on the y-axis indicate full location and abbreviations are on the x-axis. Lines indicate breaks for the main groups (*O. mykiss*, *O. m. nelsoni*, NSMO, *O. chrysogaster*, SSMO, *O. apache*, *O. gilae*, and *O. clarkii* ssp., Mexican and California hatcheries.