SPATIAL AND TEMPORAL GENETIC STRUCTURE OF WINTER-RUN STEELHEAD (*ONCORHYNCHUS MYKISS*) RETURNING TO THE MAD RIVER, CALIFORNIA

By

Steven R. Fong

A Thesis Presented to

The Faculty of Humboldt State University

In Partial Fulfillment of the Requirements for the Degree

Master of Science in Natural Resources: Fisheries

Committee Membership

Dr. Andrew Kinziger, Committee Chair

Dr. David Hankin, Committee Member

Dr. Darren Ward, Committee Member

Dr. Erin Kelly, Graduate Coordinator

July 2020

ABSTRACT

SPATIAL AND TEMPORAL GENETIC STRUCTURE OF WINTER-RUN STEELHEAD (ONCORHYNCHUS MYKISS) RETURNING TO THE MAD RIVER, CALIFORNIA

Steven R. Fong

Distinct populations of steelhead in the wild are in decline. The propagation of steelhead in hatcheries has been used to boost population numbers for recreational fisheries and for use in conservation. However, hatchery breeding practices of steelhead can result in changes in genetic structure. I investigated the genetic structure of winter-run steelhead (*Oncorhynchus mykiss*) returning to the Mad River, California, where a hatchery has been used enhance production for recreational fisheries since 1971. Genetic variability in Mad River steelhead was evaluated using 96 single nucleotide polymorphisms (SNPs) among 4203 individuals, including the Mad River and nearby locations, and spanning 44 years from 1973 to 2017. I resolved evidence that in the 1970s the Mad River contained both an indigenous population, and a population influenced by the introduction of Eel River winter-run broodstock. Even with the introduction of Eel River broodstock, contemporary Mad River steelhead (1983-2017) appear to be distinct from Eel River collections, as well as other surrounding collections (except Redwood Creek). This distinction is a consequence of the presence of a historically unique population in the Mad River, combined with the inability of the initially introduced Eel River steelhead broodstock to establish itself. Lastly, I found that contemporary Mad

River Hatchery broodstock are composed of three groups (or broodlines), defined by adult return year (1) 2009, 2012, and 2015, (2) 2010, 2013, and 2016, and (3) 2011 and 2014. Grouping in 3-year intervals is hypothesized to be a result of the predominant usage of age-3 individuals as broodstock at Mad River Hatchery.

ACKNOWLEDGEMENTS

Funding for this research project was provided by the California Department of Fish and Wildlife's (CDFW) Steelhead Fishing Report and Restoration Card program. Additional funding was provided by: The Marin Rod & Gun Club Award, The Joseph Bania Award, The Graduate Equity Fellowship, The Fisheries Founding Faculty Award, and the Golden Gate Angling Club Award. I offer you my utmost thanks and appreciation to all of these organizations.

First and foremost, I would like to thank Dr. Andrew Kinziger. Andrew has served as a wonderful mentor and has become a good friend over the course of my education at Humboldt State University. Without him I would not be the person I am today, and through his guidance and support I have achieved things in my life that I would have never thought possible. A month into starting my graduate career I was overwhelmed by the daunting prospect of it. I will never forget what he said to me… "There is only one way to eat an elephant… a bite at a time." These words helped me through my degree during the most challenging parts… one step at a time.

I was very lucky to have long time mentors serve on my graduate committee. Dr. David Hankin and Dr. Darren Ward helped me through some of the most challenging times in my life. Thank you so much for helping me navigate and find my way through graduate school.

I would also like to thank the CDFW staff in Arcata, especially Michael Sparkman, Philip Barrington, and Justin Garwood. They always helped answer my

questions and went out of their way to find the answer if they did not know it themselves (which was rare). They also provided me with any resources I needed including access to historic records, and to their Northern California Scale Archive.

My warmest thanks go to our collaborator Dr. Carlos Garza for allowing me to visit and train at the NMFS Southwest Fisheries Science Center in Santa Cruz, CA. He welcomed me to his lab, provided me with any resources I needed (including access to a large cache of steelhead genotypes), and answered any questions I had. The staff was great and treated me as if I were one of their own. They also made sure I was prepared to conduct laboratory research independently once I left the facility. I cannot thank enough Dr. Eric Anderson for allowing me to stay in his guest house, along with Ellen Campbell and Elena Correa for preparing and training me for my research, which included them helping me with the large genotyping effort.

I would also like to thank the National Science Foundation for granting HSU with the Fluidigm instrumentation needed for this project.

This journey would also not have been possible if it were not for everyone at HSU who helped me while I was making my way through graduate school. This includes all the graduate students I worked and collaborated with in the fisheries genetics lab, Rod Nakamoto from the USFS, Fisheries Faculty, Fisheries and USGS Coop administration (Katherine Moon, and Leslie Farrar), and to David Baston for allowing me to have a workspace in the CNRS Core Research Facility.

And to my family Katherine, Sonny, Richard, and Kevin Fong and Steven Chan who were with me every step of the way, and believed I could do it, thank you. Finally, to

Janett Mata who has begrudgingly let me return to grad school (twice), and still supports me along the way, thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

*Oncorhynchus mykiss (*Walbaum 1792), also known as steelhead trout (anadromous form) or rainbow trout (resident form), are one of the most commonly occurring fish in lakes, rivers, and streams on the West Coast of the United States (Moyle 2002). Many natural-origin populations of steelhead are in decline, so there has been increased effort to develop strategies for their conservation (NOAA 2006; Moore et al. 2014; Crozier et al. 2019; Schaefer et al. 2019). Steelhead support important recreational, and Native American fisheries, and limited commercial fisheries on the West Coast of North America. Hatcheries are sometimes used to boost steelhead numbers in the wild, and they do so by use of artificial propagation. The Mad River of Northern California in particular has a robust recreational fishery as a result of their hatchery, and it helps to support the neighboring local economy. The Mad River is known for having one of the best steelhead catch rates in Northern California and comprises approximately 32% of all statewide steelhead fishing trips taken in California annually (Jackson 2007; NMFS 2016).

With a well-developed strategy, hatcheries can play a vital role in the conservation of steelhead populations. Hatcheries can be used to supplement native populations and fortify recreational fisheries (Champagnon et al. 2012; Clarke et al. 2017). In the past, a common practice in hatcheries was to spawn steelhead selectively for their size, growth rate, and/or adaptive capabilities (Donaldson et al. 1957; Chilcote et al. 1986). It was also common practice for a hatchery to use out-of-basin steelhead as

hatchery broodstock when indigenous steelhead stocks were low or unavailable (CDFW 1970 - 2000; Araki et al. 2006; Johnson et al. 2010). These past practices gave little consideration to how introduction of out-of-basin fish would affect indigenous populations.

Currently, many hatcheries are either using, or are in the process of implementing, integrated programs that place an emphasis on minimizing genetic divergence between natural populations and the hatchery populations used to augment or support those natural populations. A well-designed hatchery integration program allows the natural-origin population to drive selection and fitness of the total population. This can be achieved by ensuring that the gene flow from the natural population to the hatchery population exceeds that of the hatchery population into the natural population. A hatchery without an integrated management plan may temporarily boost numbers of steelhead in a watershed, but ultimately, may reduce productivity, fitness, and the effective size of the natural population (Reisenbichler and Rubin 1999; Chilcote 2003; Araki et al. 2007; Christie et al. 2012).

In 2016 the Mad River Hatchery (MRH) implemented an official integrated management plan for steelhead (NMFS 2016). This plan prescribed the integration of, at a minimum, 50% natural-origin steelhead with a maximum of 50% hatchery-origin steelhead in a 1:1 sex ratio for their broodstock annually (NMFS 2016). Since the inception of the integrated management plan, it has been difficult to meet natural-origin steelhead targets, because natural-origin steelhead do not return to the hatchery in sufficient numbers to meet this target (Kinziger et al. 2018). To increase the numbers of

natural-origin spawners in the hatchery, a program was developed during the 2014-2015 spawning year utilizing angling stewards. The stewards collect natural-origin steelhead from various locations in the Mad River, then deliver them to the hatchery to be spawned (M. Sparkman pers. comm. 2019). Returns of natural-origin winter-run steelhead to the MRH have been a problem since the hatchery's inception due to low numbers of winter-run steelhead in the river at the time (Busby et al. 1996; NMFS 2016). This resulted in the use of out-of-basin Eel River winter-run steelhead broodstock to supplement the MRH for the first 3 years of the hatchery's operation from 1971 to 1973 (CDFW 1970 - 2000).

During the first year of the MRH's operation in January 1971 (the 1970-1971 season), a total of 301 adult winter-run steelhead (144 females, and 157 males) were taken at the Benbow Dam fish ladder, from the South Fork of the Eel River, to be used as broodstock (CDFW 1970 - 2000). A conflicting report states that eggs were taken from a different location (the Cape Horn Dam egg collecting station), but official MRH stocking records do not corroborate this information (CDFW 1970 - 2000; Busby et al. 1996). The Eel River steelhead spawned in 1971 served as the hatchery's founding winter-run broodstock (i.e. no indigenous Mad River steelhead were used as broodstock the first year) and the offspring of those steelhead were released into the Mad River as yearlings (spring of 1972) (CDFW 1970 - 2000).

For the two subsequent seasons (1971-1972, 1972-1973), a combination of Benbow Dam Eel River and Mad River natural-origin winter-run steelhead were used as broodstock. For these two seasons the MRH received few Mad River natural-origin

steelhead returns to the hatchery (1971-1972 = 42, 1972-1973 = 52). This required additional supplementation with Benbow Dam Eel River steelhead (1971-1972 = 452, 1972-1973 = 395). For the 1973-1974 spawning season, the MRH no longer supplemented their broodstock, and with few exceptions (discussed below), only used steelhead that returned to the hatchery's fish ladder (CDFW 1970 - 2000; NMFS 2016).

Since the 1970s there have been several cases where small numbers of out-of-basin stock have been (or may have been) incubated or reared at MRH and released into the Mad River. Summer-run steelhead eggs (1971 and 1973) and fingerlings (1980) of Washougal River origin and eggs (1972) and fingerlings (1973) from the Trinity Hatchery of Eel River origin were introduced into the Mad River (CDFW 1970 - 2000; NMFS 2016). During 1984 and 1985, winter-run steelhead eggs from Warm Springs Hatchery (Dry Creek origin) were introduced into the Mad River (CDFW 1970 - 2000). Finally, there was a potential introduction of steelhead from the San Lorenzo River that were reported to have been introduced in 1972 (Busby et al. 1996; Good et al. 2005), but official hatchery stocking records do not corroborate that these steelhead were released into the Mad River at any time and were only incubated, reared, and planted into nearby watersheds (CDFW 1970 - 2000).

Several studies have tried to place Mad River steelhead within the broader phylogeny of the species from throughout the Pacific Northwest of North America (Reisenbichler et al. 1992; Busby et al. 1994,1996; Bjorkstedt et al. 2005; Garza et al 2014). A transition point (or genetic shift) with reduced gene flow in coastal California steelhead has been identified in the vicinity of Humboldt Bay. Steelhead originating from

the Mad River, which is just north of Humboldt Bay, clustered with the southern collections (similar to Eel River collections) presumably as a result of the use of Eel River steelhead as broodstock at MRH in the 1970s (Garza et al. 2014).

A study of genetic structure within Mad River steelhead using variation in 14 microsatellite loci identified a temporal transition in genetic structure between the 1970s and more contemporary collections (Reneski 2011). The historical collections (taken by creel sample) of steelhead from the Mad River dating to the 1970s clustered with the Eel River whereas more contemporary collections from the 2000s were distinct from the Eel River and other collections investigated (e.g. Washougal River, San Lorenzo River, and Russian River). The temporal transition was attributed to drift within the Mad River steelhead population resulting from the use of small broodstock numbers in some years at MRH (Reneski 2011).

To elucidate potential genetic effects driven by hatchery practices in both historic and contemporary Mad River winter-run steelhead collections, I used both temporal and spatial analysis along with a comprehensive set of collections from nearby locations. Previous studies investigated either spatial or temporal relationships (Reneski 2011; Garza et al. 2014). I examined variation in 96 SNP loci (previous studies used either 14 or 15 microsatellite loci) to evaluate the effects, if any, of past hatchery management practices on the contemporary Mad River steelhead population structure. I specifically examined: (1) the impacts of using out-of-basin steelhead as broodstock at the Mad River Hatchery on genetic structure, (2) the genetic distinctiveness of historical (1970's) and contemporary Mad River steelhead, and (3) contemporary genetic structure of the Mad

River Hatchery broodstock across a 9 year period (2009 – 2017). The results of this study will assist the hatchery in managing their winter-run steelhead program.

MATERIALS AND METHODS

Sample Selection and SNP Genotyping

The dataset was composed of a total of 4203 steelhead from 47 different collections, representing a 44-year time span between 1973 to 2017, that were genotyped at 96 single nucleotide polymorphisms (SNPs) (Figure 1; Table 1). I genotyped a collection of scale samples from the Mad River. These consisted of historical 1970's collections, one 1983 collection, and contemporary creel survey collections from 1999 – 2003. These scale collections were provided by California Department of Fish and Wildlife (CDFW). These data were combined with (1) SNP genotypes from Mad River Hatchery steelhead broodstock from 2009-2017 (Kinziger et al. 2018), and (2) SNP genotypes for putative Mad River natural-origin steelhead collections from 1999-2003, three juvenile Mad River steelhead collections from 2014, and steelhead collections from outside of the Mad River collected from 2003-2014 and provided by the NOAA Southwest Fisheries Science Center (SWFSC; Table 1). The vast majority of collections in this study consisted of genotypes from winter-run steelhead, but a small subset was discovered to be of summer-run origin (Table 1). In Mad River collections from 1973-1974 and 1974-1975 there were individuals collected from September to November, indicating these were likely summer-run steelhead. Winter-run steelhead in the Mad River generally do not return to spawn until between late December to March. Use,

handling, and curation of steelhead fish scales were first approved on 31 March 2015 under the HSU IACUC permit number 14/15.F.78-E.

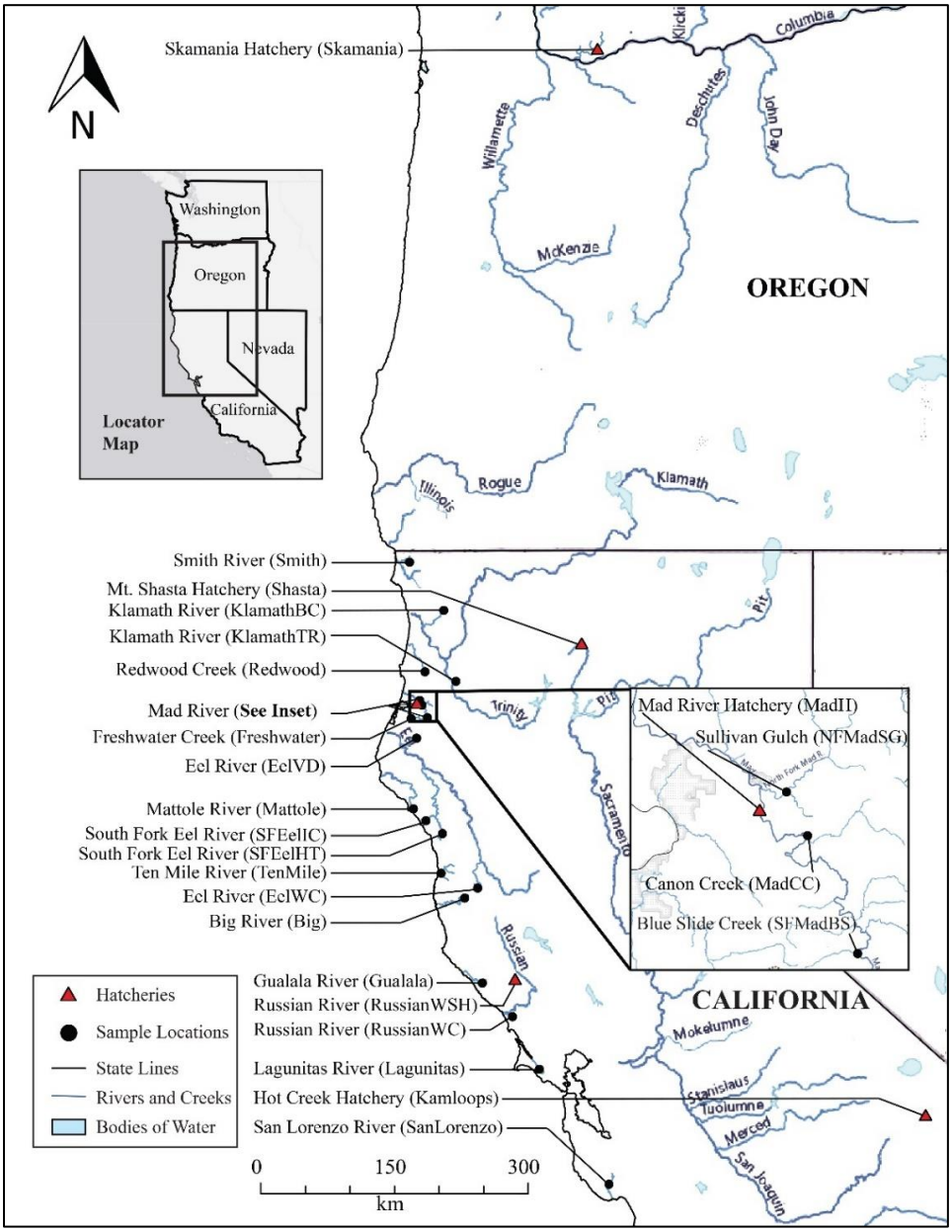

Figure 1. Geographic map of steelhead and rainbow trout collection locations taken from California and Washington in the United States. Inset of the Mad River included to detail sampling locations. All collections are labeled with their respective location, and the abbreviations (Pop. Code) used to identify them in this study. Detailed information can be found in Table 1.

Table 1. Summary of *Oncorhynchus mykiss* (steelhead) population collections and their sampling location (n = 4203). This includes watershed of origin (Origin), where samples were collected (Sample Location), population abbreviations (Pop. Code), year of collection, initial sample size (Total (N)), SNP genotype origin: "S" for DNA extracted from a scale sample, and "G" for pre-existing SNP genotype (Genetic Sample), and other identifying information such as: summer-run or winter-run, adult or juvenile, method of collection, and anadromy (Life Stage and/or Type).

Origin	Sample Location	Pop. Code	Collection Year	Total (N)	Genetic Sample	Life Stage and/or Type
Washougal River	Skamania Hatchery	Skamania	2008	45	G	Adult Summer-Run Broodstock
Hot Creek Hatchery	Hot Creek Hatchery	Kamloops	2003	47	G	Juvenile Rainbow Trout Stock (Kamloops Strain)
Mt. Shasta Hatchery	Mt. Shasta Hatchery	Shasta	2003	47	G	Juvenile Rainbow Trout Stock (Mt. Shasta Strain)
Smith River	SF Rowdy Creek	Smith	2014	75	G	Juvenile Survey
Klamath River	Trinity River	KlamathTR	2014	75	G	Juvenile Survey
Klamath River	Blue Creek	KlamathBC	2014	75	G	Juvenile Survey
Redwood Creek	Panther Creek	Redwood	2014	75	G	Juvenile Survey
Mad River	Mad River Hatchery	MadH17	2017	240	G	Adult Winter-Run Broodstock
Mad River	Mad River Hatchery	MadH16	2016	229	G	Adult Winter-Run Broodstock
Mad River	Mad River Hatchery	MadH15	2015	181	G	Adult Winter-Run Broodstock
Mad River	Mad River Hatchery	MadH14	2014	28	G	Adult Winter-Run Broodstock
Mad River	Mad River Hatchery	MadH13	2013	229	G	Adult Winter-Run Broodstock

Origin	Sample Location	Pop. Code	Collection Year	Total (N)	Genetic Sample	Life Stage and/or Type
Mad River	Mad River Hatchery	MadH12	2012	230	G	Adult Winter-Run Broodstock
Mad River	Mad River Hatchery	MadH11	2011	240	G	Adult Winter-Run Broodstock
Mad River	Mad River Hatchery	MadH09-10	2010	109	G	Adult Winter-Run Broodstock
Mad River	Mad River Hatchery	MadH08-09	2009	86	G	Adult Winter-Run Broodstock
Mad River	Mad River	MadW02-03	2002-2003	18	G	Adult Creel/Hook-line survey
Mad River	Mad River	MadW01-02	2001-2002	43	G	Adult Creel/Hook-line survey
Mad River	Mad River	MadW00-01	2000-2001	35	G	Adult Creel/Hook-line survey
Mad River	Mad River	MadW99-00	1999-2000	41	G	Adult Creel/Hook-line survey
NF Mad River	Sullivan Gulch	NFMadSG	2014	75	G	Juvenile Survey
Mad River	Canon Creek	MadCC	2014	75	G	Juvenile Survey
SF Mad River	Blue Slide Creek	SFMadBS	2014	75	G	Juvenile Survey
	Below Mad River				S	
Mad River	Hatchery	MadC02-03	2002-2003	96		Adult Creel Survey
	Below Mad River				S	
Mad River	Hatchery	MadC01-02	2001-2002	96		Adult Creel Survey

Origin	Sample Location	Pop. Code	Collection Year	Total (N)	Genetic Sample	Life Stage and/or Type
	Below Mad River				S	
Mad River	Hatchery	MadC00-01	2000-2001	96		Adult Creel Survey
	Below Mad River				S	
Mad River	Hatchery	MadC99-00	1999-2000	96		Adult Creel Survey
Mad River	Mad River	Mad83	1983	31	S	Adult 1980's Creel
Mad River	Mad River	Mad76-77	1976-1977	75	S	Adult 1970's Winter-Run Creel
Mad River	Mad River	Mad75-76	1975-1976	238	S	Adult 1970's Winter-Run Creel
Mad River	Mad River	Mad74-75	1974-1975	150	S	Adult 1970's Winter-Run Creel
Mad River	Mad River	Mad74SUM	1974	12	S	Adult 1970's Winter-Run Creel
Mad River	Mad River	Mad73-74	1973-1974	9	S	Adult 1970's Winter-Run Creel
Mad River	Mad River	Mad73SUM	1973	22	S	Adult 1970's Winter-Run Creel
Humboldt Bay	Freshwater Creek	Freshwater	2014	75	G	Juvenile Survey
Eel River	Willits Creek	EelWC	2014	46	G	Juvenile Survey
SF Eel River	Indian Creek	SFEelIC	2014	75	G	Juvenile Survey
Eel River	Van Duzen River	EelVD	2014	75	G	Juvenile Survey

Origin	Sample Location	Pop. Code	Collection Year	Total (N)	Genetic Sample	Life Stage and/or Type
SF Eel River	Hollow Tree Creek	SFEelHT	2014	75	G	Juvenile Survey
Mattole River	SF Bear Creek	Mattole	2014	75	G	Juvenile Survey
Ten Mile River	SF Ten Mile River	TenMile	2014	75	G	Juvenile Survey
Big River	NF Big River	Big	2014	75	G	Juvenile Survey
Gualala River	Fuller Creek	Gualala	2014	75	G	Juvenile Survey
Russian River	Warm Springs Hatchery	RussianWSH	2011	46	G	Adult Winter-Run Broodstock
Russian River	Willow Creek	RussianWC	2014	70	G	Juvenile Survey
Lagunitas Creek	Olema Creek	Lagunitas	2014	75	G	Juvenile Survey
San Lorenzo River	Zayante Creek	SanLorenzo	2014	72	G	Juvenile Survey

Genomic DNA was extracted from dried scales collected by the CDFW (designated as "S" in Table 1). DNeasy 96 Blood and Tissue Kits (Qiagen Inc.) were used for DNA extraction with a protocol modified for use on a Qiagen BioRobot 3000 workstation. Whole genomic DNA was pre-amplified to increase target strand concentration before SNP genotyping. Pre-amplification PCR reactions were conducted in a reaction volume of 5.4µL consisting of 2.5µL QIAGEN multiplex master mix (Catalog #206143), 1.13µL nuclease-free water, 0.15µL of TaqMan® multiplex primer pool (Applied Biosystems), and 1.6µl of diluted whole genomic DNA diluted to a ratio of 1:2. The TaqMan® multiplex primer pool uses the same primers as the SNP assay, just minus the probes (See Appendix). Pre-amplification thermal cycling conditions consisted of 95°C for 15 minutes, followed by 14 cycles of, ramp 2°C/sec to 95°C, 95°C for 15 seconds, ramp 2°C/sec to 60°C, 60°C for 4 minutes, then a final step of 10°C forever after the last cycle.

Samples were genotyped at 96 SNP loci utilizing TaqMan® SNP assays (Applied Biosystems) (Abadia-Cardoso et al. 2013; Kinziger et al. 2018; See Appendix). After PCR pre-amplification, target strands of DNA were diluted by adding 15µL of 2µM Tris buffer directly to each well of the PCR plate. Genotyping was conducted using either a Fluidigm Juno/Biomark or EP1 system (Fluidigm Corp., South San Francisco, CA) with a 96.96 Dynamic Array™ IFC. The SNP genotyping protocol can be found on the Fluidigm website (PN68000129 E1). Modifications made to the protocol for this study were optimizations to final reagent volumes.

Data Standardization and Quality Control

To ensure consistency in genotype scores generated at the SWFSC using a Fluidigm EP1 system and Humboldt State University (HSU) using a Fluidigm Juno/Biomark HD an identical set of 96 individuals were genotyped at each lab. A total of 12 of the 9216 (0.13%) scored SNP genotypes were discordant between labs (e.g. one lab called a SNP AA while the other lab called AG). Discordant genotypes were detected at *Omy_aspAT-123* (2 individuals), *Omy_R04944* (1 Individual), *SH95318-147* (5 individuals), and *SH102505-102* (4 Individuals). It was determined that discordance at loci *SH95318-147* and *SH102505-102* could be explained due to multiple clusters that made these loci hard to score accurately on a consistent basis between labs.

Prior to quality checking, the sex identification marker was removed from the dataset (the sex identification marker does not measure genetic variation). The dataset was then filtered to remove any individuals missing ≥12 loci (or approx. 13%) from their multilocus genotypes. Of the 4203 samples, 4124 remained after removal of individuals missing ≥12 loci. In addition, three loci were removed from the dataset, including two markers that were missing a considerable amount of data from certain collections, and one marker that was not common to all data sets. Thus, the final dataset consisted of 92 SNP loci and 4124 individuals. These dataset will be referred to as “The Complete” dataset.

Sibling Removal

A parentage reconstruction analysis was performed to identify and reduce the effect of large full-sibling groups from the estimates of genetic structure. Inclusion of large full-sibling groups (which can sometimes be problematic in juvenile steelhead collections) can bias estimates of linkage disequilibrium, Hardy-Weinberg equilibrium, population structure, and effective population size (Anderson and Dunham 2008; Garza et al. 2014; Rodriguez-Ramilo et al. 2014; Waples and Anderson 2017).

I conducted a parentage analysis using the software COLONY 2.0 (Wang 2004; Wang and Santure 2009) using default settings, with the exception of males and females set to polygamy and set to possible inbreeding. Full-siblings were removed following the "Yank-2 method" (Waples and Anderson 2017), which eliminates individuals within full-sibling groups that are ≥3 in size, at random, until only two siblings remain in the group. A side effect of the "Yank-2 method" is that it produces multiple datasets due to the randomized nature of the sibling removal process, but it has been found that the use of any one of these randomized datasets makes very little difference in downstream analysis (Garza et al. 2014). The dataset produced for this study using the "Yank-2 method" will be referred to as the full-siblings removed dataset or "FSR".

Genetic Diversity

Tests for conformance to Hardy-Weinberg expectations (HWE) were conducted using the software GENODIVE version 3.0 (Meirmans and Van Tienderen 2004) using

10,000 permutations of the data. Tests for HWE significance was evaluated using an uncorrected p-value ($\alpha = 0.05$) and Bonferroni correction for multiple tests ($\alpha = (0.05 / (92 \text{ SNPs} * 47 \text{ collection})) = 1.16 \times 10^{-5}$) (Rice 1989). Tests for linkage disequilibrium (LD) were performed using Markov Chain Monte Carlo using the program Genepop 4.7.0 (Raymond and Rousset 1995; Rousset 2008) using the default settings. Tests for LD significance were checked using an uncorrected ($\alpha = 0.05$) and Bonferroni corrected ($\alpha = (0.05 / 196742 \text{ tests}) = 2.54 \times 10^{-7}$) (Rice 1989) level of significance. Analyses were run on both the Complete and FSR datasets.

Estimates of observed heterozygosity (H_o), expected heterozygosity (H_e), and heterozygote deficiency (G_{IS}) were calculated using GENODIVE. Unstandardized allelic richness (A) and standardized allelic richness equalized using rarefaction to a sample size of 10 (Ar), were calculated in HP-Rare version 1.0 (Kalinoski 2004; Kalinoski 2005). Pairwise estimates of genetic differentiation (F_{ST}) between collections were calculated using GENODIVE on both the Complete and FSR dataset. To test for differences in population differentiation caused by inclusion of full-siblings, a regression was performed to compare genetic differentiation (F_{ST}) estimated in the Complete versus the FSR dataset.

Genetic Structure Analysis

Neighbor joining tree

To elucidate spatial and temporal genetic structure an unrooted neighbor-joining (N-J) tree was constructed based upon the FSR dataset using the software PHYLIP

(Felsenstein 2004). The Cavalli-Sforza (1967) method was used to estimate genetic distances between pairs of collections, with distances used to construct a N-J tree. To evaluate branch support, a bootstrap analysis was conducted (1000 replicates). Three collections had negative branch lengths (Smith, Mad73SUM, and MadW01-02) and were corrected using methods described by Kuhner and Felsenstein (1994).

Bayesian cluster analysis

The Bayesian clustering algorithm implemented in the program STRUCTURE 2.3.4 (Pritchard et al. 2000; Falush et al. 2003, 2007; Hubisz et al. 2009) was used to generate an estimate of the number of genetically distinct groups or clusters (k) in the data, and estimate the fraction of the multilocus genotype that belongs to each cluster (q). This analysis was carried out on the FSR dataset. The analysis was run assuming the data contained between 1 and 10 distinct clusters and 20 independent runs were conducted at each k. Each run was performed at 150,000 iterations (with 50,000 discarded as burn-in) using the default analysis parameters. The software STRUCTURE HARVESTER (Earl 2012) was used to summarize the log probability of the data (Pritchard et al. 2000; Falush et al. 2003, 2007; Hubisz et al. 2009) and to calculate the ad hoc statistic Δk (Evanno et al. 2005). These are two metrics that informally point to the number of clusters that best fit the data. It has been suggested that Δk can more accurately detect the most likely number of clusters in the data then the plateau of the log probability of the data. The Δk metric is based on the rate of change in the log probability of the data between successive k values (Evanno et al. 2005). The STRUCTURE output was processed using

DISTRUCT (Rosenberg 2004) to facilitate visual comparison of clusters and aid in observing shifts in genetic structure across all 47 collections of individuals.

Discriminant analysis of principal components

Lastly, I used a multivariate approach to elucidate genetic differentiation between collections. A discriminant analysis of principal components (DAPC) was performed on the FSR dataset. This multivariate approach first transforms the data using a principal component analysis (PCA), then conducts a discriminant analysis (DA). An advantage of this analysis is that it is independent of Hardy-Weinberg equilibrium and/or linkage disequilibrium assumptions. The DAPC was conducted using the *adegenet* package (Jombart 2008; Jombart & Ahned 2011) for the statistical software R version 3.6.1 (R Core Team 2019).

RESULTS

Sibling Removal

There were initially n = 4203 individuals in the dataset; after quality checking, n was reduced to 4124 individuals, the number of individuals in the Complete dataset (Table 2). For the Complete dataset, the largest group of full-siblings in the dataset consisted of 23 individuals (SFMadBS); the smallest of the full-sibling groups consisted of 2 individuals and required no removal of siblings (MadH14, MadW99-00, Mad73-74, and MadC02-03); and two collections contained only one individual with no detected siblings (MadW02-03 and RussianWSH). One collection had a large number of full-siblings removed which consisted of 112 individuals (MadH12) and 6 collections had no full-siblings removed. A total of n = 3042 individuals remained after full-sibling removal forming the FSR dataset (Table 3).

Table 2. Complete SNP dataset summary statistics. Total initial sample size (Total (N)), number of steelhead missing >12 SNP loci (#Missing >12 loci), sample size after steelhead missing >12 SNP loci are removed (Complete (N)), percent of collections not meeting Hardy-Weinberg expectations with an uncorrected $\alpha < 0.05$ (HWEU (%)), Percent of collections not meeting Hardy-Weinberg expectations using Bonferroni correction $\alpha < 1.16 \times 10^{-5}$ (HWEB (%)), percent of collections in linkage disequilibrium with an uncorrected $\alpha < 0.05$ (LDU (%)), and percent of collections in linkage disequilibrium Bonferroni corrected $\alpha < 2.54 \times 10^{-7}$ (LDB (%)).

Pop. Code	Total (N)	# Missing >12 loci	Complete (N)	HWEU (%)	HWEB (%)	LDU (%)	LDB (%)
Skamania	45	0	45	8.70	0.00	2.46	0.00
Kamloops	47	0	47	2.17	0.00	3.03	0.00
Shasta	47	0	47	2.17	0.00	3.66	0.07
Smith	75	1	74	7.61	0.00	7.45	0.02
KlamathTR	75	0	75	8.70	0.00	8.27	0.02
KlamathBC	75	0	75	4.35	0.00	5.40	0.02
Redwood	75	1	74	18.48	3.26	21.19	0.55
MadH17	240	0	240	19.57	2.17	15.05	0.19
MadH16	229	0	229	11.96	3.26	9.72	0.14
MadH15	181	0	181	13.04	0.00	10.44	0.12
MadH14	28	0	28	4.35	0.00	4.73	0.02
MadH13	229	0	229	16.30	2.17	14.88	0.17

Pop. Code	Total (N)	# Missing >12 loci	Complete (N)	HWEU (%)	HWEB (%)	LDU (%)	LDB (%)
MadH12	230	0	230	17.39	2.17	17.27	0.31
MadH11	240	0	240	17.39	0.00	15.03	0.19
MadH09-10	109	0	109	8.70	1.09	11.87	0.02
MadH08-09	86	0	86	6.52	1.09	6.90	0.05
MadW02-03	18	0	18	3.26	0.00	3.13	0.00
MadW01-02	43	1	42	6.52	0.00	8.07	0.12
MadW00-01	35	1	34	8.70	1.09	4.83	0.00
MadW99-00	41	10	31	7.61	1.09	4.68	0.00
NFMadSG	75	0	75	22.83	6.52	22.53	1.10
MadCC	75	1	74	14.13	0.00	10.46	0.12
SFMadBS	75	1	74	20.65	3.26	19.09	0.48
MadC02-03	96	1	95	11.96	1.09	6.02	0.05
MadC01-02	96	1	95	7.61	0.00	6.21	0.02
MadC00-01	96	1	95	5.43	0.00	5.38	0.05
MadC99-00	96	0	96	3.26	1.09	6.81	0.05

Pop. Code	Total (N)	# Missing >12 loci	Complete (N)	HWEU (%)	HWEB (%)	LDU (%)	LDB (%)
Mad83	31	2	29	10.87	0.00	7.53	0.00
Mad76-77	75	1	74	13.04	0.00	8.15	0.14
Mad75-76	238	45	193	29.35	7.61	14.62	1.05
Mad74-75	150	3	147	11.96	3.26	11.23	0.24
Mad74SUM	12	0	12	4.35	0.00	2.91	0.00
Mad73-74	9	0	9	2.17	0.00	1.65	0.00
Mad73SUM	22	2	20	5.43	1.09	5.95	0.00
Freshwater	75	0	75	19.57	6.52	18.59	0.53
EelWC	46	1	45	6.52	0.00	8.67	0.02
SFEelIC	75	1	74	13.04	1.09	23.58	0.93
EelVD	75	1	74	6.52	0.00	6.62	0.00
SFEelHT	75	1	74	4.35	0.00	9.34	0.05
Mattole	75	0	75	11.96	1.09	19.47	0.43
TenMile	75	0	75	8.70	3.26	13.66	0.12
Big	75	0	75	11.96	0.00	9.41	0.02

Pop. Code	Total (N)	# Missing >12 loci	Complete (N)	HWEU (%)	HWEB (%)	LDU (%)	LDB (%)
Gualala	75	0	75	16.30	3.26	22.91	0.81
RussianWSH	46	0	46	3.26	0.00	4.92	0.00
RussianWC	70	1	69	20.65	5.43	23.75	1.41
Lagunitas	75	0	75	11.96	0.00	12.85	0.17
SanLorenzo	72	2	70	8.70	0.00	12.92	0.07

Table 3. Full Siblings Removed (FSR) SNP dataset summary statistics. Sample size after steelhead missing >12 SNP loci are removed (Complete (N)), number of siblings removed from the Complete dataset (Siblings Removed), sample size after siblings were removed (FSR (N)), percent of collections not meeting Hardy-Weinberg expectations with an uncorrected $\alpha < 0.05$ (HWEU (%)), percent of collections not meeting Hardy-Weinberg expectations using Bonferroni correction $\alpha < 1.16 \times 10^{-5}$ (HWEB (%)), percent of collections in linkage disequilibrium with an uncorrected $\alpha < 0.05$ (LDU (%)), percent of collections in linkage disequilibrium Bonferroni corrected $\alpha < 2.54 \times 10^{-7}$ (LDB (%)), measure of heterozygote deficiency in the population (G_{IS}), expected heterozygosity (*He*), observed heterozygosity (*Ho*), average number of successfully genotyped SNP loci (Ave L), average allelic richness (Ave A), average allelic richness using rarefaction standardized to a sample size of 10 (Ave Ar), and average differentiation calculated across the remaining 46 collections (Ave F_{ST}).

Pop. Code	Complete (N)	Siblings Removed	FSR (N)	HWEU (%)	HWEB (%)	LDU (%)	LDB (%)	G_{IS}	*He*	*Ho*	Ave L	Ave Ar	Ave F_{ST}
Skamania	45	30	15	2.17	0.00	0.88	0.00	0.037	0.22	0.21	91.07	1.58	0.217
Kamloops	47	22	25	2.17	0.00	2.20	0.00	0.007	0.23	0.23	91.76	1.63	0.223
Shasta	47	28	19	0.00	0.00	2.75	0.00	-0.002	0.32	0.32	91.11	1.76	0.209
Smith	74	9	65	6.52	0.00	6.14	0.02	0.017	0.33	0.32	91.63	1.81	0.109
KlamathTR	75	13	62	9.78	0.00	5.66	0.02	0.007	0.31	0.30	91.56	1.79	0.139
KlamathBC	75	2	73	4.35	0.00	4.87	0.02	0.026	0.30	0.29	91.74	1.77	0.123
Redwood	74	19	55	6.52	0.00	8.98	0.05	0.043	0.36	0.35	91.25	1.87	0.048
MadH17	240	91	149	6.52	1.09	7.53	0.05	0.014	0.33	0.33	91.68	1.81	0.053
MadH16	229	71	158	11.96	1.09	6.47	0.07	0.000	0.34	0.34	91.49	1.82	0.053
MadH15	181	44	137	6.52	0.00	7.57	0.07	-0.001	0.34	0.34	90.74	1.82	0.052

Pop. Code	Complete (N)	Siblings Removed	FSR (N)	HWEU (%)	HWEB (%)	LDU (%)	LDB (%)	G_{IS}	*He*	*Ho*	Ave L	Ave Ar	Ave F_{ST}
MadH14	28	0	28	4.35	0.00	4.71	0.02	0.006	0.33	0.33	91.29	1.81	0.053
MadH13	229	95	134	16.30	1.09	7.48	0.05	0.003	0.34	0.34	91.30	1.82	0.052
MadH12	230	112	118	9.78	0.00	7.00	0.05	0.002	0.34	0.34	91.41	1.82	0.053
MadH11	240	65	175	9.78	0.00	9.15	0.12	0.003	0.33	0.33	91.37	1.81	0.054
MadH09-10	109	24	85	6.52	1.09	8.19	0.02	0.026	0.34	0.33	91.67	1.83	0.050
MadH08-09	86	5	81	5.43	1.09	6.12	0.02	0.009	0.34	0.33	91.70	1.82	0.049
MadW02-03	18	0	18	3.26	0.00	3.15	0.00	0.054	0.36	0.34	91.72	1.87	0.049
MadW01-02	42	2	40	4.35	0.00	7.88	0.07	0.074	0.37	0.34	91.58	1.88	0.056
MadW00-01	34	1	33	8.70	1.09	4.52	0.00	0.078	0.37	0.34	91.48	1.88	0.040
MadW99-00	31	0	31	7.61	1.09	4.71	0.00	0.042	0.34	0.33	91.48	1.84	0.044
NFMadSG	75	53	22	3.26	0.00	4.99	0.00	-0.075	0.32	0.34	91.73	1.76	0.066
MadCC	74	14	60	9.78	0.00	7.02	0.02	0.007	0.34	0.34	91.58	1.83	0.048
SFMadBS	74	30	44	4.35	0.00	7.21	0.00	0.002	0.36	0.36	91.80	1.86	0.041
MadC02-03	95	0	95	11.96	1.09	6.02	0.07	0.034	0.34	0.33	91.25	1.82	0.046
MadC01-02	95	3	92	7.61	0.00	6.14	0.02	0.008	0.33	0.33	91.47	1.81	0.048

Pop. Code	Complete (N)	Siblings Removed	FSR (N)	HWEU (%)	HWEB (%)	LDU (%)	LDB (%)	G_{IS}	*He*	*Ho*	Ave L	Ave Ar	Ave F_{ST}
MadC00-01	95	4	91	5.43	0.00	5.35	0.02	-0.008	0.34	0.34	91.16	1.81	0.049
MadC99-00	96	6	90	3.26	1.09	6.33	0.07	0.015	0.34	0.33	91.23	1.83	0.048
Mad83	29	6	23	8.70	0.00	5.11	0.00	0.009	0.35	0.35	90.13	1.86	0.055
Mad76-77	74	3	71	14.13	0.00	6.83	0.07	0.061	0.35	0.33	90.58	1.86	0.047
Mad75-76	193	8	185	29.35	5.43	13.28	0.69	0.101	0.36	0.33	89.09	1.88	0.041
Mad74-75	147	15	132	10.87	2.17	7.88	0.07	0.027	0.36	0.35	89.70	1.87	0.041
Mad74SUM	12	1	11	3.26	0.00	2.32	0.00	0.172	0.35	0.29	91.18	1.86	0.061
Mad73-74	9	0	9	2.17	0.00	1.65	0.00	0.077	0.36	0.33	91.33	1.86	0.044
Mad73SUM	20	2	18	6.52	0.00	4.92	0.00	0.162	0.37	0.31	90.11	1.88	0.061
Freshwater	75	43	32	5.43	0.00	6.33	0.02	-0.032	0.32	0.34	91.66	1.77	0.085
EelWC	45	7	38	3.26	0.00	6.19	0.00	0.023	0.35	0.34	91.82	1.85	0.067
SFEelIC	74	32	42	3.26	0.00	6.86	0.00	0.051	0.34	0.32	91.55	1.83	0.065
EelVD	74	4	70	3.26	0.00	5.92	0.00	0.017	0.35	0.34	91.59	1.85	0.052
SFEelHT	74	8	66	2.17	0.00	8.03	0.02	0.004	0.34	0.34	91.83	1.83	0.072
Mattole	75	35	40	3.26	0.00	7.12	0.02	-0.024	0.34	0.35	91.78	1.83	0.086

Pop. Code	Complete (N)	Siblings Removed	FSR (N)	HWEU (%)	HWEB (%)	LDU (%)	LDB (%)	G_{IS}	*He*	*Ho*	Ave L	Ave Ar	Ave F_{ST}
TenMile	75	28	47	4.35	2.17	6.86	0.00	-0.001	0.34	0.34	91.49	1.84	0.083
Big	75	10	65	9.78	0.00	7.29	0.00	0.010	0.36	0.36	91.57	1.87	0.083
Gualala	75	49	26	4.35	1.09	5.40	0.00	-0.002	0.35	0.35	91.81	1.86	0.103
RussianWSH	46	0	46	3.26	0.00	4.83	0.00	0.025	0.37	0.36	90.83	1.88	0.052
RussianWC	69	49	20	3.26	0.00	4.83	0.02	-0.100	0.36	0.40	91.60	1.85	0.137
Lagunitas	75	18	57	7.61	0.00	8.29	0.02	0.013	0.37	0.36	91.68	1.88	0.071
SanLorenzo	70	21	49	6.52	0.00	7.21	0.05	-0.032	0.36	0.37	91.94	1.87	0.122

Genetic Diversity

For both the Complete and FSR datasets, Hardy-Weinberg expectations (HWE) and linkage disequilibrium (LD) were represented as the proportion of loci exhibiting significant departures from expectations. For HWE (Bonferroni corrected) the Complete dataset exhibited variable but relatively low levels of departures at each loci for 22 out of 47 collections (range: 0.00 – 7.61; Table 2), while the FSR dataset had 13 of the 47 collections exhibiting variable but relatively low levels of departures at each loci (range: 0.00 – 5.43; Table 3). For LD (Bonferroni corrected) the Complete dataset exhibited variable but relatively low levels of LD at each loci in 36 of 47 collections (range: 0.00 – 1.41; Table 2), while the FSR dataset exhibited variable but relatively low levels of LD at each loci in 28 of 47 collections (range: 0.00 – 0.69; Table 3). With a few exceptions, the FSR dataset exhibited a reduction in HWE departures (19%) and LD (17%) when compared to the Complete dataset (Table 2 and Table 3). Removal of full-siblings demonstrated the effectiveness of the "Yank-2 method" at reducing the proportion of departures from HWE and amount of LD, but it did so at the cost of a reduction in sample size (a loss of 1082 out of 4124 samples).

Average pairwise genetic differentiation (F_{ST}) were calculated among the 46 collections (Table 3). Comparison of pairwise genetic differentiation in the Complete and FSR datasets indicated that datasets were highly correlated ($R^2 = 0.9875$, $P<0.0001$) (Figure 2). Unless otherwise noted the remaining analysis were of the FSR dataset.

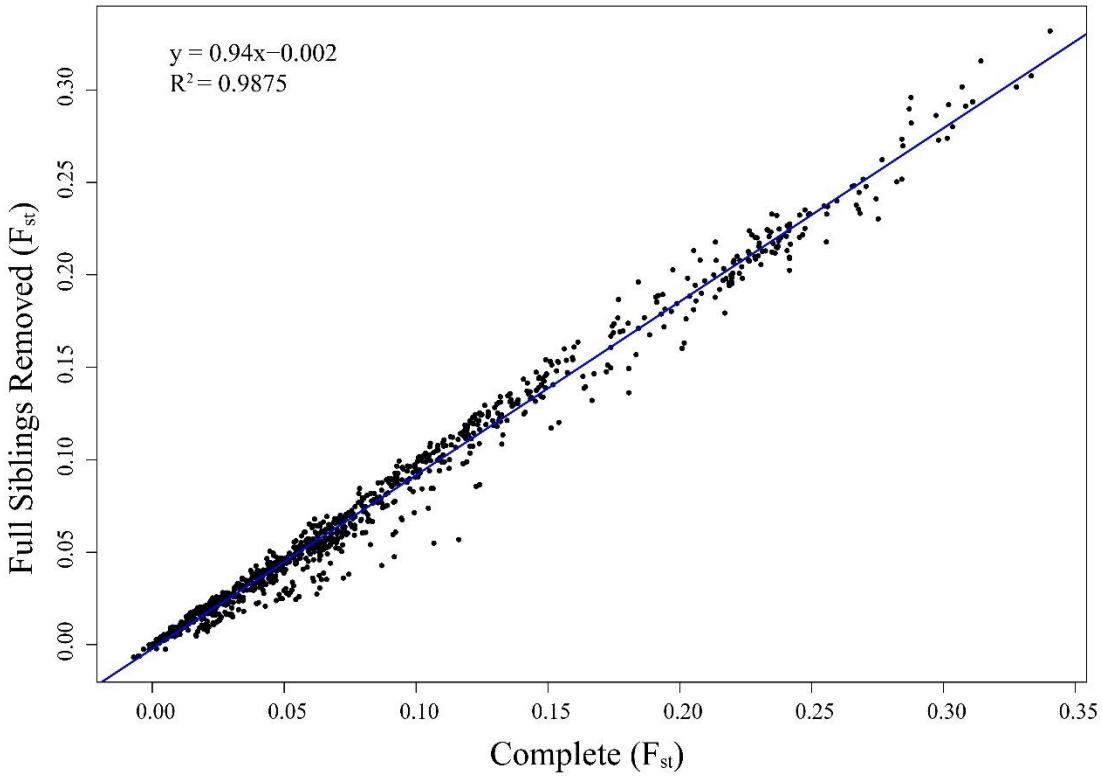

Figure 2. Regression of F_{ST} values for both the Complete and Full Siblings Removed (FSR) datasets to compare genetic distances. Values were highly correlated (R^2 = 0.9875, P<0.0001) indicating full-sibling removal had little effect on genetic differentiation.

Expected heterozygosity (H_e) ranged from 0.32 to 0.37 in the Mad River collections, and from 0.22 to 0.37 across all collections (Table 3). Observed heterozygosity (H_o) ranged from 0.29 to 0.36 in the Mad River collections, and from 0.21 to 0.40 across all collections (Table 3). Average inbreeding coefficient (G_{IS}) ranged from -0.075 to 0.172 in the Mad River collections, and from -0.100 to 0.172 across all collections. There was evidence that the 1970's collections exhibited a notable deficiency of heterozygotes and departures from HWE (Waples 2015). All six 1970's collections exhibited $H_o < H_e$, and only one collection (Mad74-75) did not have a notably elevated $G_{IS} >0$ (but the value was >0) (Table 3).

Mean number of successfully genotyped loci (out of 92) for all individuals in a collection (Ave. L) ranged from 89.09 to 91.80 in the Mad River collections, and 89.09 to 91.94 across all collections (Table 3). Allelic richness using rarefaction standardized to a sample size of 10 (Ave. A_r) ranged from 1.76 to 1.88 in the Mad River collections, and 1.58 to 1.88 across all collections (Table 3). Minor allele frequency for the Mad River collections consisted of 25 loci <0.15, 27 between 0.15 and 0.30, and 40 > 0.30. Minor allele frequency for all collections were 24 SNP loci <0.15, 27 between 0.15 and 0.30, and 41 > 0.30.

Genetic Structure

Neighbor joining tree

In the N-J tree the clustering pattern generally coincided with the geographic location of each collection (Figure 1), following a linear arrangement of collections from

north to south along the California coast (except Lagunitas, RussianWSH, Freshwater and Redwood; Figure 3). The N-J tree places the historical Mad River winter-run collections (1973-1977) in close relation to Eel River steelhead (Figure 3b). A notable exception to this result was that the Mad River summer-run steelhead collections from the 1970s (Mad73SUM and Mad74SUM) are placed separately and are well-supported (bootstrap value = 93%) within the clade of the northern steelhead collections (Figure 3b). The Mad River winter-run steelhead from the 1970s summer-run collections, the 1983 collection, and the 1999 - 2017 contemporary collections appear as unique, but weakly diverged, from the 1970s Mad River winter-run steelhead, the Eel River, and all other surrounding watersheds (except Redwood Creek).

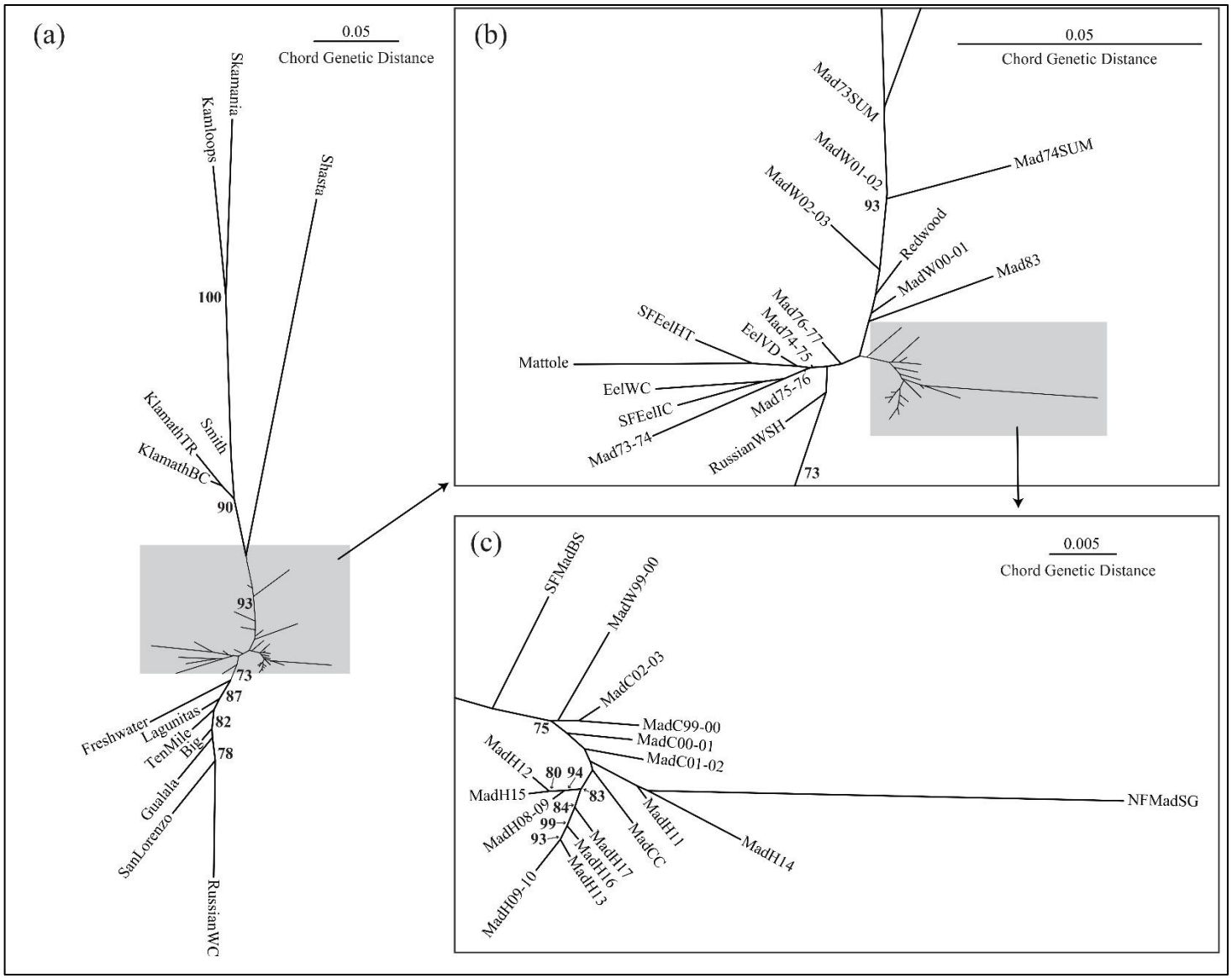
(a)
0.05
Chord Genetic Distance
Skamania
Kamloops
Shasta
100
KlamathTR
Smith
KlamathBC
90
93
73
87
82
78
Freshwater
Lagunitas
TenMile
Big
Gualala
SanLorenzo
RussianWC
(b)
0.05
Chord Genetic Distance
Mad73SUM
MadW01-02
Mad74SUM
93
MadW02-03
Redwood
MadW00-01
Mad83
Mad76-77
Mad74-75
EelVD
SFEelHT
Mattole
EelWC
SFEelIC
Mad73-74
Mad75-76
RussianWSH
73
(c)
0.005
Chord Genetic Distance
SFMadBS
MadW99-00
MadC02-03
75
MadC99-00
MadC00-01
MadC01-02
MadH12
80
94
MadH15
83
MadH08-09
84
99
93
MadH09-10
MadH17
MadH16
MadH13
MadH11
MadCC
MadH14
NFMadSG

Figure 3. Neighbor-joining tree constructed using Full Siblings Removed (FSR) dataset based upon Cavalli-Sforza chord distances. Bootstrap support was calculated using 1000 replicates with bootstrap support >70% shown. (a) The full N-J tree organized from north (top) to south (bottom), (b) Inset of Eel River, Mad River, and associated collections, and (c) inset of contemporary Mad River collections including juvenile surveys from 2014 (SFMadBS, NFMadSG, MadCC), creel surveys from 1999-2003 (MadC), and hatchery broodstock from 2009-2017 (MadH).

Within the contemporary collections from the Mad River Hatchery broodstock (and to a small degree in the creel collections) there was evidence of three separate groups (or broodlines) of adult steelhead organized by a three-year return pattern (Figure 3c). This pattern exists as steelhead returning in 2009, 2012, and 2015 (bootstrap value = 94%), 2010, 2013, and 2016 (bootstrap value = 99%), and a third group 2011 and 2014 (not bootstrap supported; Figure 3c). The third group should hypothetically include collections from 2017, but this exception is described in the Discussion. There was one example of this potential three-year return pattern existing within the creel collections in which MadC99-00 and Mad02-03 are closely related (Figure 3c).

An additional N-J tree analysis was conducted using a truncated FSR dataset including only Eel River and Mad River collections. There were little to no differences in topology and estimated bootstrap support values in comparison to the full FSR dataset. Since these results did offer any new information they were not included in the results.

Bayesian structure analysis

The STRUCTURE analysis for $k = 2$ and $k = 3$ both converged to a single result for all 20 out of 20 runs, while $k = 4$ converged into two separate results (e.g. 14 out of 20 runs converged to one result, while 6 out of 20 runs converged to different result; Figure 4). For $k = 2$, all 20 out of 20 runs detected a single abrupt structural shift in q values between Redwood Creek (Redwood) and the Klamath River (KlamathBC; Figure 4). The result for $k = 3$, for all 20 out of 20 runs, detected the same structural shift in q values at $k = 2$, but additionally detected an abrupt structural shift in q values between Mad River Creel 1999 – 2000 (MadC99-00) and Mad River Creel 1983 (Mad83; Figure 4). Finally,

at $k = 4$, the same structural shifts in q values for both $k = 2$ and $k = 3$ were detected in 14 out of 20 runs, but a new shift in q values was detected for 6 of the 20 runs. In those 6 runs, a shift in q values was detected between the Mattole River (Mattole) and Ten Mile River (TenMile; Figure 4).

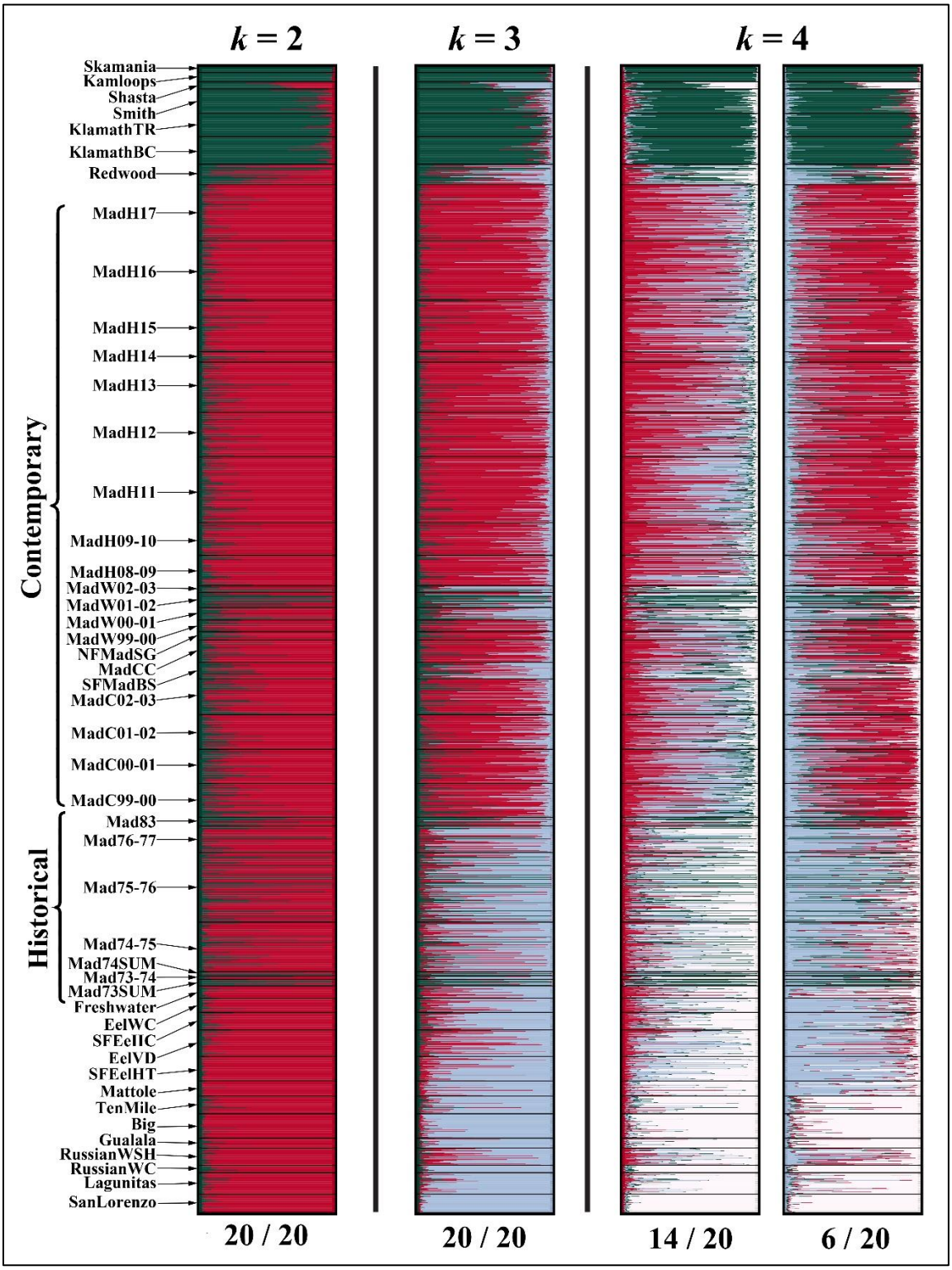
k = 2
k = 3
k = 4
Skamania
Kamloops
Shasta
Smith
KlamathTR
KlamathBC
Redwood
MadH17
MadH16
MadH15
MadH14
MadH13
MadH12
MadH11
MadH09-10
MadH08-09
MadW02-03
MadW01-02
MadW00-01
MadW99-00
NFMadSG
MadCC
SFMadBS
MadC02-03
MadC01-02
MadC00-01
MadC99-00
Mad83
Mad76-77
Mad75-76
Mad74-75
Mad74SUM
Mad73-74
Mad73SUM
Freshwater
EelWC
SFEelIC
EelVD
SFEelHT
Mattole
TenMile
Big
Gualala
RussianWSH
RussianWC
Lagunitas
SanLorenzo
Contemporary
Historical
20 / 20
20 / 20
14 / 20
6 / 20

Figure 4. Bayesian cluster analysis for the Full Siblings Removed (FSR) dataset. The value of *k* represents a posterior estimation of population structure and forms clusters (*k*) based on the fraction of the individuals multilocus genotype belonging to each cluster (*q*) and was estimated from $k = 1$ to $k = 10$, with each collection separated by a black line. Estimation of *k* was run independently for 20 iterations. Values for $k = 2$ to $k = 3$ converged to one result in all 20 out of 20 runs, while $k = 4$ converged into two different results. Collection locations are listed from north (top) to south (bottom), with Mad River collections designated as contemporary, or historical, and are sorted sequentially by year.

Based on the above results, $k = 3$ was selected as most representative of the population structure of these data for two reasons. (1) The shifts seen in $k = 3$ were all seen in $k = 4$, with the exception of the single shift seen in only 6 runs. This new shift began to split individuals between clusters suggesting that k was being overestimated (Figure 4). (2) The log probability of the STRUCTURE data arrived at a plateau at about $k = 3$, and the ad hoc statistic Δk indicated the strongest level of structure at $k = 2$ clusters ($\Delta k = 43.7$) and $k = 3$ clusters ($\Delta k = 25.6$), but there was little evidence for $k = 4$ ($\Delta k < 1.5$; Table 4).

The STRUCTURE analysis for $k = 3$ displayed 3 distinct groups (1) a southern group (SanLorenzo to Mad76-77), (2) a contemporary Mad River group (MadC99-00 to MadH17), and (3) a northern group (KlamathBC to Skamania; Figure 4). These clusters at $k = 3$ corroborate the results and approximate geographic population placements observed in the N-J tree (Figure 3).

Table 4. Results from the program STRUCTURE HARVESTER used to infer which number of clusters (*k*) best fits the data using the log probability (Mean LnP(*k*)) of the data and the ad hoc statistic (Δ*k*). The strongest level of inferred structure can be determined when the log probability reaches a plateau near the k that best fits the data, and when Δ*k* no longer detects a change in the log probabilities (e.g. Δ*k* after *k* = 4 stabilizes at about 1.27).

k	Reps	Mean LnP(*k*)	Stdev LnP(*k*)	Δ*k*
1	20	-297054.58	0.40	NA
2	20	-291099.94	68.73	43.67
3	20	-288146.62	90.42	25.57
4	20	-287505.33	206.33	1.27
5	20	-286601.85	324.03	1.28
6	20	-286111.83	307.61	1.19
7	20	-285987.91	393.18	1.05
8	20	-285450.67	617.82	0.73
9	20	-285361.27	506.76	0.60
10	20	-284968.65	465.41	NA

A second and third analysis was conducted, separate from Figure 4, using STRUCTURE on a truncated FSR dataset containing (1) only collections within the Mad River and (2) only Mad River broodstock (2009-2017). This was done to tease out any fine-scale structure that could have been obstructed by the use of the comparative collections (both outgroups, and within basin). There was no difference in genetic structure detected under either scenario. Since these results did offer any new information they were not included in the results.

Discriminant analysis of principal components

The discriminant analysis of principal components (DAPC), for all FSR sample collections, resolved a linear arrangement of steelhead collections, from north (top) to south (bottom; Figure 5). This was a similar arrangement to the neighbor-joining tree (Figure 3) and the STRUCTURE analysis (Figure 4). Collections such as the historical Mad River summer-run steelhead (Mad73SUM and Mad74SUM) and Redwood Creek (Redwood) show placement between Mad River contemporary collections and northern collections (Figure 5). Also, historical Mad River winter-run collections from the 1970's cluster with Eel River collections.

An additional DAPC was created, separate from Figure 5, was created for just Mad River and Eel River collections but did not result in a difference in population structure or an increase or change of genetic distance between groups. Since these results did offer any new information they were not included in the results.


DA eigenvalues
Skamania
Kamloops
KlamathTR
KlamathBC
Smith
Mad74SUM
MadW01−02
Redwood
Shasta
MadW02−03
NFMadSG
SFMadBS
Mad76−77
MRH17
EelWC
RussianWSH
Freshwater
Mattole
TenMile
Gualala
Big
RussianWC

Figure 5. Discriminant analysis of principal components (DAPC) of the Full Siblings Removed (FSR) dataset for all collections. Collections clustered geographically from the north (top) to the south (bottom) pattern, with label colors representing inferred degrees of relatedness. Colors and collections correspond with the Bayesian cluster analysis (Figure 4) for $k = 3$.

DISCUSSION

Genetic Relationships of Steelhead Returning to the Mad River

Prior genetic analyses of Mad River steelhead have generally concluded that early use of Eel River broodstock in the hatchery changed contemporary Mad River steelhead genetic structure (Bjorkstedt et al. 2005; Spence et al. 2008; Garza et al 2014). Garza et al. (2014) hypothesized that the vicinity of Humboldt Bay may have been a transition point between two steelhead groups and that the transfer of Eel River steelhead into the Mad River obscured a precise transition point. This was based on the observation that steelhead from Humboldt Bay (Freshwater Creek) were found to be genetically similar to Klamath River steelhead, whereas Mad River steelhead (from the first basin north of Humboldt Bay) were genetically closer to Eel River steelhead (Garza 2014).

The use of both spatial and temporal steelhead collections in my analysis (as opposed to one or the other) seems to have helped resolve some confusion over Mad River genetic structure. Finding placement of contemporary Mad River steelhead in relation to other collections along a spatial scale, while remaining in the context of a historic temporal scale, has allowed us to see how the genetic structure of Mad River steelhead has evolved and changed over time. In the 1970's Mad River steelhead genetic structure was split between winter-run steelhead grouping with Eel River steelhead, and summer-run steelhead grouping with northern collections (Figure 3). But contemporary Mad River winter-run steelhead collections form a well-supported group that is spatially

located north of Humboldt Bay, and is genetically closer to steelhead from Redwood Creek than from the Eel River (Figure 3b).

Also, in contrast to Garza et al. 2014, I found that Humboldt Bay steelhead (Freshwater Creek) were closer in relation to collections south of the Eel River (Figure 3). A potential reason for this discordant result between studies may be related to Coastal California steelhead following an isolation by distance (IBD) model (Wright 1943; Garza et al. 2014). The IBD model states that geographical distance between populations is correlated with genetic distance between populations. In steelhead, IBD suggests that watersheds that are close to one another should have a higher rate of gene flow and migration then distant watersheds. This was demonstrated in the neighbor-joining tree (Figure 3) and the DAPC analysis (Figure 5). The close proximity of these collections around Humboldt Bay may offer insight to potential temporal allelic frequency shifts as a result of steelhead migration and geneflow described by the IBD model.

1970's Mad River Steelhead

My results support the hypothesis that two genetically distinct groups of winter-run steelhead occurred in the Mad River in the 1970's: an indigenous Mad River stock and an introduced Eel River stock. This conclusion is supported by the clustering of the 1970's winter-run and summer-run collections from the Mad River. While the 1970's winter-run collections were genetically similar to the Eel River collection, the 1970's summer-run steelhead collections (Mad73SUM and Mad74SUM) cluster with collections closer to the north, exhibiting a distinction from the 1970's Mad River winter-run

samples in both tree-based and Bayesian clustering analysis (Figures 3b and 4). The summer-run collections likely share the same neutral gene pool as the indigenous Mad River winter-run steelhead of the 1970's, as winter-run and summer-run steelhead are generally identical at neutral loci (Hess et al. 2016). Thus, the summer-run collections should be genetically similar to winter-run Mad River steelhead that were present prior to operations of the Mad River Hatchery.

Additional evidence for the existence of mixed stocks owes to the presence of a Wahlund effect among the 1970's Mad River collections (Table 3). A Wahlund effect is characterized as a collection of individuals having a $H_o<H_e$, an elevated G_{IS} that is >0, and exhibits disruptions in HWE which could indicate a deficiency in heterozygotes in the total population. This deficiency is generally caused by the presence of two distinct subpopulations within a single population, that have differing allele frequencies (Wahlund 1928; Allendorf et al. 2008). The genetic findings supporting the presence of two genetically distinct groups of steelhead, residing in the Mad River during the 1970s, and is consistent with Reneski's (2011) analysis of microsatellite loci. That study found there was large-scale use of Eel River steelhead at the Mad River Hatchery in the early 1970's, while natural-origin steelhead were still present in the river. This is also supported by official hatchery stocking records from the time (CDFW 1970 - 2000).

Contemporary Mad River Steelhead

Despite the use of several out-of-basin stocks at the Mad River Hatchery, no one out-of-basin stock has had a large enough influence to affect the contemporary Mad

River winter-run steelhead population structure. The predominant use of Eel River broodstock at the Mad River Hatchery is shown in the N-J tree with all 1970's winter-run collections grouping with the Eel River (Figure 3b). The Bayesian cluster analysis is consistent with this pattern as all historical 1970's winter-run steelhead samples cluster with Eel River collections (Figure 4). However, my analysis shows that contemporary collections of steelhead from the Mad River do not cluster with the Eel River collections (Figures 3, 4, and 5) suggesting that the Eel River and Mad River steelhead have not been genetically homogenized.

The lack of impact of out-of-basin stocks on within basin genetic structure has also been reported in Klamath River Chinook salmon, where out-of-basin stocks were introduced, but failed to establish themselves and had no lasting effect on genetic structuring (Kinziger et al. 2013). In contrast if repeated introductions occur over an extended period of time, genetic homogenization of stocks may result (Williamson and May 2005). Interestingly, the Mad River collection from 1983 (Mad83) may serve as a window that shows the 1970's Mad River winter-run steelhead transitioning and returning back to its pre-hatchery genetic structure. This is based on the collection's location on the N-J tree, and DAPC (Figure 3 and 5), in addition to individual cluster assignments in the Bayesian cluster analysis (Figure 4) appearing north of the contemporary collections.

The collection from 1983, along with contemporary Mad River steelhead collections (1999-2017) appear as unique, but weakly diverged, from all surrounding watersheds (except Redwood Creek; Figure 3b). The distinction of Mad River steelhead

is most evident in the Bayesian cluster analysis where the contemporary Mad River collection is identified as a unique cluster in $k = 3$ (Figure 4). However, my analysis also suggests genetic divergence among the contemporary Mad River collections. In particular the contemporary Mad River Hatchery broodstock collections from 2009-2017 form a well-supported branch in the tree-based analysis (Figure 3c). I hypothesize that this grouping is a result of hatchery management practices, such as the use of a small effective number of breeders, and closure of broodstock to natural-origin immigration (up until 2014) (Reneski 2011). These factors, combined with the predominance of hatchery-produced steelhead in the Mad River, likely account for the contemporary, within basin changes in allele frequency.

Formation of 3-year Broodlines

Within Mad River Hatchery steelhead broodstock, I found evidence for the formation of three genetically similar temporal groups (or broodlines). These broodlines are represented as genetically similar groups in the data that return to spawn in a three-year pattern (e.g. a 2009, 2012, 2015 broodline, a 2010, 2013, and 2016 broodline, and a 2011 and 2014 broodline) (Figure 3c). This grouping by broodline has been identified in the genetic analysis of other hatcheries involving other types of Pacific salmonids (Van Doornik et al. 2002; Smith et al. 2015; Garza, unpublished). I hypothesize that grouping by broodlines at Mad River Hatchery is due to the use of a high proportion of age-3 steelhead as broodstock and limited use of age-2 and/or age-4+ steelhead.

An example of broodlines on a broader scale is evident in the NFMadSG collection (collected in 2014), which are closely related to the 2014 hatchery broodstock (MadH14) (Figure 3c). There could be two explanations for this connection (1) steelhead were released as offspring from the 2014 hatchery broodstock then sampled at NFMadSG, or (2) they are non-hatchery origin steelhead that share the same hatchery broodline as a result of many years of hatchery influence on the total population (Garza, unpublished; Figure 3c).

Evidence of broodline formation is also evident among the Mad River Hatchery steelhead broodstock collected in 2017. Due to a lawsuit filed by the Environmental Protection Information Center (EPIC) in 2013, the Mad River Hatchery was forced to shut down production for seven weeks during the 2014 spawning season. When the lawsuit was settled the hatchery resumed operations on 5 February 2014. By this time there were only 5 weeks remaining in the 2014 spawning season. Due to limiting factors caused by the newly implemented 50% natural-origin broodstock integration in a 1:1 sex ratio (as part of the settlement) only 21 females were spawned during 2014 (NMFS 2016). Low production levels presumably led to a shortage of sexually mature age-3 steelhead returning to the Mad River Hatchery in 3 years to spawn during the 2017 season.

In the 2017 spawning season, hatchery broodstock age distribution was skewed from its normal average distribution of age classes (2017: 19.9% age-2, 32.8% age-3, and 47.3% age-4 steelhead; typical year: 6.4% age-2, 70.3% age-3, 22.9% age-4, and 0.4% age-5) (Kinziger et al. 2018). This shortage of returning age-3 steelhead in 2017 partially

dismantled the 2011, 2014 broodline clustering on the tree because there were a greater number of age-4 steelhead within the total composition of the 2017 broodstock (47.3%). This resulted in 2017 broodstock (MRH17) clustering with the 2010, 2013, 2016 broodline, and not with 2011 and 2014 broodline as would be expected if a higher proportion of age-3 steelhead were used in the 2017 broodstock (Figure 3c).

Too little is known about broodlines to know whether mitigating management measures are required for steelhead or the Mad River Hatchery steelhead program. It is not known whether broodlines are a naturally occurring phenomenon or if hatchery mating practices have artificially induced broodline structuring. In coho salmon (*Oncorhynchus kisutch*), broodline structuring was deemed unnatural because genetic differentiation among broodlines exceeded geographic structure, which is atypical for Pacific salmon (Van Doornik et al. 2002, Smith et al. 2015). In the Quilcene Hatchery the exclusion of age-2 coho salmon broodstock, for nearly a century, lead to the creation of three populations from one origin population (Smith et al. 2015). So, at least for coho salmon in this system, this highlighted the importance of understanding age-2 fish contribution to the population, and how age-at-maturity shapes the populations genetic structure (Smith et al. 2015).

If age-2 steelhead were to be incorporated into the Mad River Hatcheries annual broodstock important questions would need to be answered: (1) How many age-2 steelhead should be incorporated into annual broodstock? and (2) What number of age-2 steelhead would represent natural contribution? Initially the Mad River Hatchery was advised against using age-2 steelhead in their broodstock and that they should be totally

excluded (NMFS 2016). Then shortly after, the National Marine Fisheries Service (NMFS) suggested the Mad River Hatchery incorporate age-2 steelhead into their broodstock at about 1-2% annually, but the actual percentage of age-2 steelhead that represent a “natural” contribution is currently unknown and is being investigated (NMFS 2016).

LITERATURE CITED

Abadía-Cardoso, A., A.J. Clemento, and J.C. Garza. 2011. Discovery and characterization of single-nucleotide polymorphisms in steelhead/rainbow trout, Oncorhynchus mykiss. Molecular Ecology Resources 11:31-49.

Abadía-Cardoso, A., E.C. Anderson, D.E. Pearse, and J.C. Garza. 2013. Large-scale parentage analysis reveals reproductive patterns and heritability of spawn timing in a hatchery population of steelhead (*Oncorhynchus mykiss*). Molecular Ecology 22(18):4733-4746.

Aguilar, A., and J.C. Garza. 2008. Isolation of 15 single nucleotide polymorphisms from coastal steelhead, Oncorhynchus mykiss (Salmonidae). Molecular Ecology Resources 8(3):659-662.

Allendorf, F.W., G. Luikart, and S.N. Aitken. 2008. Conservation and the genetics of populations. Blackwell Publishing, Massachusetts.

Anderson, E.C., and K.K. Dunham. 2008. The influence of family groups on inferences made with the program Structure. Molecular ecology resources 8(6):1219-1229.

Araki, H., B. Cooper, and M.S. Blouin. 2007. Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. Science 318(5847):100-103.

Araki, H., W.R. Ardren, E. Olsen, B. Cooper, and M.S. Blouin. 2006. Reproductive success of captive-bred steelhead trout in the wild: evaluation of three hatchery programs in the Hood River. Conservation Biology 21(1):181-190.

Bjorkstedt, E.P., B.C. Spence, J.C. Garza, D.G. Hankin, D. Fuller, W.E. Jones, J.J. Smith, and R. Macedo. 2005. An analysis of historical population structure for evolutionary significant units of Chinook salmon, coho salmon, and steelhead in the North-Central California Coast Recovery Domain. US Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service [Northwest Fisheries Science Center]. NOAA-TM-NMFS-SWFSC-382.

Brunelli, J.P., K.J. Wertzler, K. Sundin, and G.H. Thorgaard. 2008. Y-specific sequences and polymorphisms in rainbow trout and Chinook salmon. Genome 51(9):739-748.

Busby, P.J., T.C. Wainwright, and R.S. Waples. 1994. Status review for Klamath Mountains Province steelhead. US Department of Commerce, National Oceanic

and Atmospheric Administration, National Marine Fisheries Service [Northwest Fisheries Science Center].

Busby, P.J., T.C. Wainwright, G.J. Bryant, L.J. Lierheimer, R.S. Waples, F.W. Waknitz, and I.V. Lagomarsino. 1996. Status review of west coast steelhead from Washington, Idaho, Oregon, and California. U.S. Department of Commerce, NOAA Technical Memorandum, NMFS-NWFSC-27.

Campbell, N.R., and S.R. Narum. 2009. Quantitative PCR assessment of microsatellite and SNP genotyping with variable quality DNA extracts. Conservation Genetics 10(3):779-784.

Cavalli-Sforza, L. L., and A. F. W. Edwards. 1967. Phylogenetic analysis: models and estimation procedures. American Journal of Human Genetics 19:233-257.

CDFW (California Department of Fish and Wildlife). 1970-2000. Mad River Salmon and Steelhead Hatchery, Blue Lake, CA years 1970-2000, 31 reports. California Department of Fish and Game Hatchery-Data Reports 73-12, 74-1, 74-2, 75-1, 76-7, 78-16, 78-17, 79-3, 80-10, 82-1, 82-4, 82-27, 83-13, 86-02, 87-4, 87-10, 88-3, 90-2, 90-14, 91-9, 92-10, 94-3, 1994, 1994, 94-3, 1995, 1996, 1997, 1998, 1999, 2000, Arcata, CA.

Champagnon, J., J. Elmberg, M. Guillemain, M. Gauthier-Clerc, and J.D. Lebreton. 2012. Conspecifics can be aliens too: a review of effects of restocking practices in vertebrates. Journal for Nature Conservation 20(4):231-241.

Chilcote, M.W. 2003. Relationship between natural productivity and the frequency of wild fish in mixed spawning populations of wild and hatchery steelhead (*Oncorhynchus mykiss*). Canadian Journal of Fisheries and Aquatic Sciences 60(9):1057-1067.

Chilcote, M.W., S.A. Leider, and J.J. Loch. 1986. Differential reproductive success of hatchery and wild summer-run steelhead under natural conditions. Transactions of the American Fisheries Society 115(5):726-735.

Christie, M.R., M.L. Marine, R.A. French, R.S. Waples, and M.S. Blouin. 2012. Effective size of a wild salmonid population is greatly reduced by hatchery supplementation. Heredity 109(4):254-260.

Clarke, L.R., M.W. Flesher, W.J. Knox, and R.W. Carmichael. 2017. Increased harvest of anadromous hatchery steelhead, *Oncorhynchus mykiss* (Walbaum), through return timing manipulation. Fisheries Management and Ecology 24(4):274-282.

Crozier, L.G., M.M. McClure, T. Beechie, S.J. Bograd, D.A. Boughton, M. Carr, T.D. Cooney, J.B. Dunham, C.M. Greene, M.A. Haltuch, and E.L. Hazen. 2019. Climate vulnerability assessment for Pacific salmon and steelhead in the California Current Large Marine Ecosystem. PloS one 14(7).

Donaldson, L.R., and P.R. Olson. 1957. Development of rainbow trout brood stock by selective breeding. Transactions of the American Fisheries Society 85(1):93-101.

Earl, D.A. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation genetics resources 4(2):359-361. doi: 10.1007/s12686-011-9548-7

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular ecology 14(8):2611-2620.

Falush, D., M. Stephens, and J.K Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164(4):1567-1587.

Falush, D., M. Stephens, and J.K. Pritchard. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular ecology notes 7(4):574-578.

Felsenstein, J. 2004. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences University of Washington, Seattle.

Garza, J.C. 2012. Population Genetic Structure of Coho Salmon in the Klamath River. Unpublished data.

Garza, J.C., E.A. Gilbert-Horvath, B.C. Spence, T.H. Williams, H. Fish, S.A. Gough, J.H. Anderson, D. Hamm, and E.C. Anderson. 2014. Population structure of steelhead in coastal California. Transactions of the American Fisheries Society 143(1):134-152.

Good, T.P., R.S. Waples, and P.B. Adams. 2005. Updated status of federally listed ESUs of West Coast salmon and steelhead. U.S. Department of Commerce, NOAA Technical Memorandum, NMFS-NWFSC-66.

Hess, J.E., J.S. Zendt, A.R. Matala, and S.R. Narum. 2016. Genetic basis of adult migration timing in anadromous steelhead discovered through multivariate association testing. Proceedings of the Royal Society B: Biological Sciences 283(1830):2015-3064.

Hubisz, M.J., D. Falush, M. Stephens, and J.K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. Molecular ecology resources 9(5):1322-1332.

Jackson, T.A., and N. Manji. 2007. California steelhead fishing report-restoration card: a report to the legislature. State of California, The Resources Agency, Department of Fish and Game.

Johnson, S.L., J.H. Power, D.R. Wilson, and J. Ray. 2010. A comparison of the survival and migratory behavior of hatchery-reared and naturally reared steelhead smolts in the Alsea River and estuary, Oregon, using acoustic telemetry. North American Journal of Fisheries Management 30(1):55-71.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24(11):1403-1405.

Jombart, T., and I. Ahmed. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27(21):3070-3071.

Kalinowski, S.T. 2004. Counting alleles with rarefaction: private alleles and hierarchical sampling designs. Conservation genetics 5(4):539-543.

Kalinowski S.T. 2005. HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. Molecular Ecology Notes 5:187-189

Kinziger, A.P., M. Hellmair, D.G. Hankin, and J.C. Garza. 2013. Contemporary population structure in Klamath River basin Chinook salmon revealed by analysis of microsatellite genetic data. Transactions of the American Fisheries Society 142(5):1347-1357.

Kinziger, A.P., S.R. Fong, J.C. Garza. 2018. Application of parentage-based tagging for monitoring steelhead returning to Mad River Hatchery, California. California Department of Fish and Wildlife.

Kuhner, M.K. and J. Felsenstein. 1994. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. Molecular biology and evolution 11(3):459-468.

Meirmans, P.G., and P.H. Van Tienderen. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Molecular ecology notes 4(4):792-794.

Moore, J.W., J.D. Yeakel, D. Peard, J. Lough, and M. Beere. 2014. Life-history diversity and its importance to population stability and persistence of a migratory fish:

steelhead in two large North American watersheds. Journal of Animal Ecology, 83(5):1035-1046.

Moyle, P.B. 2002. Inland fishes of California: revised and expanded. University of California Press, Berkeley.

NOAA (National Oceanic Atmospheric Administration). 2006. Endangered and threatened species: final listing determinations for 10 distinct population segments of west coast steelhead. Federal Register. 50:223,224(5 January 2006):833-862.

NMFS (National Marine Fisheries Service). 2016. Hatchery and genetic management plan for Mad River Hatchery winter-run steelhead. National Oceanic and Atmospheric Administration, National Marine Fisheries Service.

Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155(2):945-959.

R Core Team 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2 updated to v3.3): population genetics software for exact tests and ecumenicism. Journal of Heredity 86:248-249.

Reisenbichler, R.R., J.D. McIntyre, M.F. Solazzi, and S.W. Landino. 1992. Genetic variation in steelhead of Oregon and northern California. Transactions of the American Fisheries Society 121(2):158-169.

Reisenbichler, R.R., and S.P. Rubin. 1999. Genetic changes from artificial propagation of Pacific salmon affect the productivity and viability of supplemented populations. ICES Journal of marine Science 56(4):459-466.

Reneski, M.R. 2011. Temporal genetic analysis of steelhead (*Oncorhynchus mykiss*) reveals hatchery-induced drift in captivity. Master's thesis. Humboldt State University, Arcata, California.

Rice, W.R. 1989. Analyzing tables of statistical tests. Evolution 43(1):223-225.

Rodríguez-Ramilo, S.T., M.A. Toro, J. Wang, and J. Fernández. 2014. Improving the inference of population genetic structure in the presence of related individuals. Genetics research 96.

Rosenberg, N. A. 2004. Distruct: a program for the graphical display of population structure. Molecular Ecology Notes 4:137-138.

Rousset, F. 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Molecular Ecology Resources 8:103-106.

Schaefer, R.A., S.L. Gallagher, and C.D. Chamberlain. 2019. Distribution and Abundance of California Central Valley steelhead/Rainbow Trout and Late-fall Chinook Salmon Redds in Clear Creek, Winter 2015 to Spring 2016. U.S. Fish and Wildlife Service, Red Bluff Fish and Wildlife Office, Red Bluff California.

Smith, C.T., J. Baumsteiger, W.R. Ardren, Y. Dettlaff, D.K. Hawkins, and D.M. Van Doornik. 2015. Eliminating variation in age at spawning leads to genetic divergence within a single Salmon population. Journal of Fish and Wildlife Management 6(1):4-18.

Spence, B.C., E.P. Bjorkstedt, J.C. Garza, J.J. Smith, D.G. Hankin, D.W. Fuller, W.E. Jones, R. Macedo, T.H. Williams, and E. Mora. 2008. A framework for assessing the viability of threatened and endangered salmon and steelhead in the North-Central California Coast Recovery Domain. U.S. Department of Commerce, NOAA Technical Memorandum, NOAA-TM-NMFS-SWFSC-423.

Van Doornik, D.M., M.J. Ford, and D.J. Teel. 2002. Patterns of temporal genetic variation in coho salmon: estimates of the effective proportion of 2-year-olds in natural and hatchery populations. Transactions of the American Fisheries Society 131(6):1007-1019.

Wahlund, R.S. 1928. The combination of populations and the appearance of correlation examined from the standpoint of the study of heredity. Hereditas 11:65-106.

Wang, J. 2004. Sibship reconstruction from genetic data with typing errors. Genetics 166:1963–1979.

Wang, J., and A. W. Santure. 2009. Parentage and sibship inference from multilocus genotype data under polygamy. Genetics 181:1579–1594.

Waples, R.S. 2015. Testing for Hardy–Weinberg proportions: have we lost the plot?. Journal of heredity 106(1):1-19.

Waples, R.S., and E.C. Anderson. 2017. Purging putative siblings from population genetic data sets: a cautionary view. Molecular Ecology 26(5):1211-1224.

Williamson, K.S. and B. May. 2005. Homogenization of fall-run Chinook salmon gene pools in the Central Valley of California, USA. North American Journal of Fisheries Management 25(3):993-1009.

Wright, S. 1943. Isolation by distance. Genetics 28(2):114.

APPENDIX

Assay ID	Reference
OMGH1PROM1-SNP1	Abadía-Cardoso et al. 2011
Omy_AldA	Aguilar & Garza 2008
Omy_arp-630	Campbell et al. 2009
Omy_aspAT-123	Campbell et al. 2009
Omy_COX1-221	Campbell et al. 2009
Omy_g12-82	J. DeKoning unpublished
Omy_gh-475	Campbell et al. 2009
Omy_gsdf-291	J. DeKoning unpublished
Omy_mapK3-103	N. Campbell unpublished
Omy_mcsf-371	J. DeKoning unpublished
Omy_nramp-146	Campbell et al. 2009
Omy_Ogo4-304	Campbell et al. 2009
OMY_PEPA-INT6	Aguilar & Garza 2008
ONMYCRBF_1-SNP1	Aguilar & Garza 2008
SEX_ID	Brunelli et al. 2008
SH100771-63	Abadía-Cardoso et al. 2011
SH100974-386	Abadía-Cardoso et al. 2011
SH101554-306	Abadía-Cardoso et al. 2011
SH101770-410	Abadía-Cardoso et al. 2011
SH101832-195	Abadía-Cardoso et al. 2011
SH101993-189	Abadía-Cardoso et al. 2011

Assay ID	Reference
SH102420-634	Abadía-Cardoso et al. 2011
SH102505-102	Abadía-Cardoso et al. 2011
SH102510-682	Abadía-Cardoso et al. 2011
SH102867-443	Abadía-Cardoso et al. 2011
SH103350-395	Abadía-Cardoso et al. 2011
SH103577-379	Abadía-Cardoso et al. 2011
SH103705-558	Abadía-Cardoso et al. 2011
SH104519-624	Abadía-Cardoso et al. 2011
SH105075-162	Abadía-Cardoso et al. 2011
SH105105-448	Abadía-Cardoso et al. 2011
SH105115-367	Abadía-Cardoso et al. 2011
SH105385-406	Abadía-Cardoso et al. 2011
SH105386-347	Abadía-Cardoso et al. 2011
SH105714-265	Abadía-Cardoso et al. 2011
SH106172-332	Abadía-Cardoso et al. 2011
SH106313-445	Abadía-Cardoso et al. 2011
SH107074-217	Abadía-Cardoso et al. 2011
SH107285-69	Abadía-Cardoso et al. 2011
SH108735-311	Abadía-Cardoso et al. 2011
SH109243-222	Abadía-Cardoso et al. 2011
SH109525-403	Abadía-Cardoso et al. 2011
SH109651-445	Abadía-Cardoso et al. 2011
SH109693-461	Abadía-Cardoso et al. 2011

Assay ID	Reference
SH109874-148	Abadía-Cardoso et al. 2011
SH110064-419	Abadía-Cardoso et al. 2011
SH110078-294	Abadía-Cardoso et al. 2011
SH110201-359	Abadía-Cardoso et al. 2011
SH110362-585	Abadía-Cardoso et al. 2011
SH110689-148	Abadía-Cardoso et al. 2011
SH111666-301	Abadía-Cardoso et al. 2011
SH112208-328	Abadía-Cardoso et al. 2011
SH112301-202	Abadía-Cardoso et al. 2011
SH112820-82	Abadía-Cardoso et al. 2011
SH113109-205	Abadía-Cardoso et al. 2011
SH113128-73	Abadía-Cardoso et al. 2011
SH114315-438	Abadía-Cardoso et al. 2011
SH114448-87	Abadía-Cardoso et al. 2011
SH114587-480	Abadía-Cardoso et al. 2011
SH114976-223	Abadía-Cardoso et al. 2011
SH115987-812	Abadía-Cardoso et al. 2011
SH116733-349	Abadía-Cardoso et al. 2011
SH117259-96	Abadía-Cardoso et al. 2011
SH117286-374	Abadía-Cardoso et al. 2011
SH117370-400	Abadía-Cardoso et al. 2011
SH117540-259	Abadía-Cardoso et al. 2011
SH117815-81	Abadía-Cardoso et al. 2011

Assay ID	Reference
SH118175-396	Abadía-Cardoso et al. 2011
SH118654-91	Abadía-Cardoso et al. 2011
SH118938-341	Abadía-Cardoso et al. 2011
SH119108-357	Abadía-Cardoso et al. 2011
SH119892-365	Abadía-Cardoso et al. 2011
SH120255-332	Abadía-Cardoso et al. 2011
SH120950-569	Abadía-Cardoso et al. 2011
SH121006-131	Abadía-Cardoso et al. 2011
SH123044-128	Abadía-Cardoso et al. 2011
SH125998-61	Abadía-Cardoso et al. 2011
SH127236-583	Abadía-Cardoso et al. 2011
SH127510-920	Abadía-Cardoso et al. 2011
SH127645-308	Abadía-Cardoso et al. 2011
SH128851-273	Abadía-Cardoso et al. 2011
SH128996-481	Abadía-Cardoso et al. 2011
SH129870-756	Abadía-Cardoso et al. 2011
SH130524-160	Abadía-Cardoso et al. 2011
SH130720-100	Abadía-Cardoso et al. 2011
SH131460-646	Abadía-Cardoso et al. 2011
SH131965-120	Abadía-Cardoso et al. 2011
SH95318-147	Abadía-Cardoso et al. 2011
SH95489-423	Abadía-Cardoso et al. 2011
SH96222-125	Abadía-Cardoso et al. 2011

Assay ID	Reference
SH97077-73	Abadía-Cardoso et al. 2011
SH97954-618	Abadía-Cardoso et al. 2011
SH98188-405	Abadía-Cardoso et al. 2011
SH98409-549	Abadía-Cardoso et al. 2011
SH98683-165	Abadía-Cardoso et al. 2011
SH99300-202	Abadía-Cardoso et al. 2011