UC Santa Cruz

UC Santa Cruz Electronic Theses and Dissertations

Title

Genetic Pedigree Inference of Coho Salmon: A Powerful Tool for Guiding the Management of an ESA-Listed Species

Permalink

https://escholarship.org/uc/item/866637qj

Author

Starks, Hilary

Publication Date

2014

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

GENETIC PEDIGREE INFERENCE OF COHO SALMON: A POWERFUL TOOL FOR GUIDING THE MANAGEMENT OF AN ESA-LISTED SPECIES

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

MASTER OF SCIENCE

in

OCEAN SCIENCES

by

Hilary A. Starks

June 2014

The Thesis of Hilary A. Starks
is approved:

Dr. John Carlos Garza, Chair

Dr. Raphael Kudela

Dr. Grant Pogson

Dean Tyrus Miller
Vice Provost and Dean of Graduate Studies

Copyright © by

Hilary A. Starks

2014

Table of Contents

List of Figures

List of Tables

Abstract

Genetic pedigree inference of coho salmon: a powerful tool for guiding the management of an ESA-listed species

by

Hilary A. Starks

Coho salmon (*Oncorhynchus kisutch*) is a species of Pacific salmon comprising an important component of the commercial and recreational fishing industry. However, over much of its native range, coho salmon population numbers have declined drastically and, as such, the species is facing extinction. For this reason, strides must be made to better manage and conserve remaining populations in order to ensure that this species will be present for future generations. This thesis provides an in depth evaluation of a population of coho salmon using pedigree inference combined with novel molecular techniques, in order to address several biological questions pertaining to the conservation of this species. In Chapter one, I describe the discovery, characterization and development of a large number of single nucleotide polymorphisms (SNPs) specific to coho salmon for the purpose of future biological inference. In Chapter two, I use these SNPs to reconstruct pedigrees, by way of intergenerational genetic tagging, in an ESA-listed hatchery population of coho salmon located in the Klamath River. Intergenerational genetic tagging involves genotyping parental individuals and using their genotypes as genetic tags that are recovered when the offspring of the parental generation are genotyped with the same markers. In this study, nine consecutive years of broodstock samples were

genotyped with 96 SNPs, and tag retrieval was possible for three full cohorts in the offspring generations, revealing large proportions of age-two males returning to spawn at the hatchery, as well as the presence of large families distributed across relatively few parent pairs. Additionally, I performed an in depth evaluation of relatedness (Rxy) in this population by comparing SNP-generated pairwise Rxy values to those generated by microsatellites, in order to test and validate SNP utility for estimates of relatedness. The results of this analysis showed that mean relatedness values generated from both marker types can be skewed by the presence of a large number of half siblings sired by one two-year old male. However, when this individual's offspring were accounted for, SNPs performed as well or better than microsatellites at estimating relatedness among individuals of known pedigree. The information provided by these tag recoveries will help us to better understand the effects of hatchery practices on hatchery spawned fish as well as the genetic effects of hatcheries on natural populations, which are important matters in regard to an ESA-listed species.

To my grandparents, who loved the sea.

Acknowledgments

I am extremely grateful to my MS advisor, Carlos Garza, for giving me the opportunity to pursue this degree. Without his support, guidance and patience this would not have been possible.

I'd also like to thank my committee members, Raphael Kudela and Grant Pogson, for taking the time to read and comment on my thesis, especially when it felt as if it would never be completed. Their patience and and flexibility has been a blessing!

I am also especially thankful to the entire Molecular Ecology team at the Santa Cruz Lab. Over the years I have gained invaluable skills and knowledge from Elizabeth Gilbert-Horvath, Devon Pearse and Eric Anderson. I want to especially thank Anthony Clemento for his continuous help (and patience) throughout the years with all things involving r, salmon or simply life. It has also been both a pleasure and a privilege to work with Alicia Abadía-Cardoso, Martha Arciniega, Eric Crandall, Vanessa Apkenas, Cassie Columbus and Diana Baetscher, all of whom I consider to be awesome coworkers as well as good friends.

Lastly, I'd like to thank my parents and my sisters, whose confidence in my abilities and unwavering encouragement, has kept me going during this experience!

Introduction

Coho salmon (*Oncorhynchus kisutch*) are distributed along the Northern Pacific Rim, ranging from Eastern Russia and around the Bering Sea, to Alaska and the North American coast, extending as far south as central California (Sandercock 1991). Recent importance has been placed on coho salmon populations within California and Oregon due to dramatic declines in population abundance over the last thirty years (Brown *et al.* 1994). Historically, the commercial and recreational coho salmon fishery has represented an important industry both economically and culturally. Presently native populations of the species have declined dramatically due to over-fishing, habitat degradation involving industrial and agricultural activities, hatchery domestication selection, disease, and the introduction of anthropogenic barriers to migration (Håstein and Lindstad 1991; Clemento *et al.* 2009; Hallett *et al.* 2012). Evolutionarily Significant Units (ESU) have been established based on the degree of genetic population structure and isolation differentiating salmonid populations (Waples 1991) and of the seven ESUs established for coho salmon within California, Oregon and Washington, five have been listed as either endangered (Central California Coast), threatened (Southern Oregon Northern California Coast; Oregon Coast; Lower Columbia River) or a species of con-

cern (Pacific-Puget Sound/Strait of Georgia) under the Endangered Species Act (ESA; Fed Reg 1995, 1997, 2005).

Molecular methods have been employed throughout the years in order to study salmonid biology and ecology. Early genetic analyses utilized allozymes, or protein polymorphisms, to discriminate salmonid populations geographically (Milner *et al.* 1985; Tessier *et al.* 1995). However, technological advancements in the field of population genetics have given rise to many more molecular markers with which to study salmonids, some of which include amplified fragment length polymorphisms (AFLPs), mitochondrial markers (mtDNA), minisatellites and microsatellites (Beacham *et al.* 1996; Smith *et al.* 2001; Flannery *et al.* 2007). Microsatellites have been used extensively for the analysis of the genetic diversity of natural populations in the last two decades due to their high variability and power to resolve population structure. Most specifically they have been recognized in being able to provide evaluations of fine-scale population structure in salmonids (Banks *et al.* 2000). Beacham *et al.* (2006) show that 90% accuracy of assignment to populations was achieved in samples of 75 individuals using 13 microsatellite loci. However, microsatellites have drawbacks in that they have high rates of mutation and a high incidence of homoplasy (Narum *et al.* 2008), as well as inconsistencies in allele calling across different laboratories due to differences in automation (Vignal *et al.* 2002).

As such, many researchers are now turning to single nucleotide polymorphisms (SNPs) for population genetic studies (Morin *et al.* 2004). SNPs are nucleotide variants that can be found at particular locations throughout an organism's genome. SNPs have

relatively low mutation rates, can occur within both genomic coding and noncoding regions, and represent the most abundant sequence variation in eukaryotic genomes (Wang *et al.* 1998; Brumfield *et al.* 2003; Smith et al. 2006). Additionally, SNPs are amenable to high-throughput genotyping and portability of data between laboratories. As such, SNPs have the potential for a wide array of applications in wildlife conservation and management.

In Chapter one, I describe the ascertainment, characterization and development of a large number of SNPs for coho salmon (*O. kisutch*), for the purpose of monitoring coho salmon populations as well as studying ecological interactions within and among populations. In Chapter two, I demonstrate the power of SNPs for intergenerational genetic tagging and pedigree reconstruction in a hatchery population of coho salmon. This effort provides a powerful and informative means of estimating relative reproductive success, effective population size, population structure and inbreeding in an ESA listed species.

Chapter 1

Discovery and characterization of single nucleotide polymorphisms in coho salmon, *Oncorhynchus kisutch*

1.1 Introduction

One of the most important ways of monitoring depressed fish and wildlife populations is through population genetics. Since the advent of the polymerase chain reaction (PCR), the development of molecular markers for the detection of genetic diversity both within and among populations has served to be extremely useful for the purpose of fish and wildlife management, particularly for monitoring salmonid stocks. A broad spectrum of molecular markers have been developed as a result, including restriction fragment length polymorphisms, randomly amplified polymorphic DNA, mitochondrial DNA, minisatellites, and microsatellites (Beacham *et al.* 1996, Smith *et al.* 2001). The

popularity of each type of marker has varied over time, but, for more than a decade, microsatellites have endured widespread use by fishery biologists (Narum *et al.* 2004; Aguilar and Garza 2006; Pearse *et al.* 2007; Beacham *et al.* 2011). Microsatellite loci are highly variable and as such offer significant statistical power for evaluations of fine-scale population structure, pedigree analyses, and genetic stock identification (GSI) in salmonids (Banks *et al.* 2000, Beacham *et al.* 2011). Microsatellites however are prone to high rates of mutation and genotyping error, as well as a high incidence of homoplasy (Narum *et al.* 2008). In addition, microsatellite data are often difficult to compare across laboratories. Discrepancies in allele size calling occur when different instrumentation is used to generate data due to differences in the chemistry and software capabilities provided by each platform, even when the same microsatellite loci are employed (Vignal *et al.* 2002, Seeb *et al.* 2007). This in turn makes the standardization of data between laboratories a costly and time consuming process, and has lead researchers to investigate other resources for genetic analysis.

As such, many labs are turning to SNPs for genetic analyses. SNPs are single nucleotide variants found in sequences of DNA. Because SNPs are ususally biallelic, larger numbers of loci are required in order to achieve analytical power similar to that of microsatellite loci (Anderson and Garza 2006, Narum *et al.* 2008). SNPs are the most abundant polymorphisms in vertebrate genomes (coding and non-coding regions), with a SNP present every 100-500bp on average (Vignal *et al.* 2002). An additional benefit to SNPs is that their genotypes are unambiguous so they do not require extensive and costly standardization efforts between laboratories, as SNP genotyping is based

on evaluating the actual nucleotides (A, C, G or T) rather than fragment length (i.e. microsatellites; Templin *et al.* 2005).

Coho salmon (*Oncorhynchus kisutch*) is one of the seven species of Pacific salmon native to the North Pacific Rim. The natural distribution of coho salmon extends from northern Japan to coastal North America, reaching as far south as central California (Sandercock 1991). Coho salmon are also stocked in the Great Lakes, where natural spawning has been documented in some Great Lake tributaries, representing one of the few successful self-sustaining introductions for this species (Behnke 2002).

Coho salmon have historically served as a prominent fishery constituent both in the recreational and commercial industry. Over-fishing, freshwater habitat alteration, changing ocean conditions (Bradford and Irvine 2000), disease (Bartholomew and Foote 2010), and hatchery domestication selection (Ford *et al.* 2006) have all been implicated however, in recent declines and even extinctions of certain populations within the species' range (Nehlsen *et al.* 1991, Small *et al.* 1998). As a result, many populations of coho salmon have been listed as endangered or threatened under the federal Endangered Species Act (Weitkamp *et al.* 1995), with the most impacted populations residing in California, the southern most reaches of the species' distribution.

Although coho salmon face high risks of extinction, only a relatively small number of SNP assays have been developed for this species (Smith *et al.* 2005, Smith *et al.* 2006, Campbell and Narum 2011). The Gene Index Project (compbio.dfci.harvard.edu /tgi/) is a database housing a large number of expressed sequence tags (ESTs) from many species, including rainbow trout, *O. mykiss*. These EST sequences have proven

to be very useful for evaluating variation in other Pacific salmonids (Smith *et al.* 2005, Abadía-Cardoso *et al.* 2011, Clemento *et al.* 2011).

In this study, the discovery, development and characterization of 91 SNP assays for coho salmon and one species diagnostic assay designed to genetically differentiate coho and Chinook salmon are described. I took advantange of an existing EST database from which 275 primer sets had been designed for functional genome regions and that produced PCR products in Chinook salmon and steelhead. These primer sets were then used to sequence an ascertainment sample of 24 geographically and phenotypically diverse coho salmon. The resulting assays were then validated by genotyping 470 coho individuals from 10 populations within the species' native range, the details of which are described below.

1.2 Methods

1.2.1 Ascertainment and PCR

An ascertainment panel was created using 24 coho salmon individuals sampled from five geographically and phylogenetically diverse populations within the species' North American range. The majority of these individuals represented prominent populations within northern California, with six fish from Scott Creek, six from the Russian River, and four from the Klamath River. The remaining samples consisted of four individuals from the Nehalem River in Oregon and four from Gastineau Hatchery in Alaska. For the California and Alaska populations, DNeasy 96 tissue kits (QIAGEN Inc.) were

used to extract DNA from fin tissue on a QIAGEN BioRobot 3000. Aliquots of DNA from the Oregon samples were extracted and provided to us by collaborators. Except for the Oregon individuals, all samples were previously genotyped with microsatellites to confirm high DNA quality.

Originally, 480 *O. mykiss* ESTs were randomly targeted from the online Gene Index database (http://compbio.dfci.harvard.edu/tgi/, accessed on December 8, 2006) for rainbow trout. For each locus, oligonucleotide primers were designed using *primer*3 v. 0.4.0 (Rozen and Skaletsky 2000). In order to generate genomic DNA fragments no larger than 1000 bp in length even if containing intronic regions, primers were constructed to target 400-500 bp EST segments. Based on prior amplification in steelhead trout and Chinook salmon (Abadía-Cardoso *et al.* 2011, Clemento *et al.* 2011), 275 of these primer sets were then selected for SNP ascertainment in coho salmon.

PCR was performed using a 15μL reaction volume with the following conditions: 1.5μL of 10X PCR buffer II (Applied Biosystems, Inc.), 0.9μL of 25 mM $MgCl_2$, 1μL of 2.5 mM dNTPs, 1μL of 5 mM primers (forward and reverse), 6.55μL of deionized water, 0.05μL of AmpliTaq DNA polymerase, and 4μL of genomic DNA. Thermal cycling conditions were modified from a step-down protocol and began with an initial denaturation of 95°C for 5 min, followed by 95°C for 3 min, 60°C for 2 min, 72°C for 1 min, repeated 13 times with a 1°C decrease in anneal temperature (60-48°C) each cycle, then 11 cycles of 95°C for 30 s, 48°C for 30 s, 72°C for 1 min, and 9 cycles of 95°C for 30 s, 48°C for 30 s, 72°C for 1 min (+10 s/cycle), with a final 5 min extension at 72°C. PCR products were visualized by gel electrophoresis in 2% agarose.

1.2.2 Sequencing and SNP Assay Development

PCR products were sequenced if a locus displayed a single band in agarose. For loci that exhibited such positive amplification, PCR products were first purified using an EXO-SAP protocol as follows: 5μL of PCR product, 0.15μL of Exonuclease I (20U/μL), 1μL of shrimp alkaline phosphatase (1 U/μL), 0.5μL of 10x buffer, and 3.35μL of deionized water, incubated at 37°C for 60 min and then 80°C for 20 min with a final cool down to 4°C. Clean PCR products were then sequenced on both the forward and reverse strands using the BigDye Terminator v. 3.1 Cycle Sequencing kit (Applied Biosystems Inc.) with standard conditions. The sequencing reaction products then underwent a final purification step using 6% Sephadex columns and were visualized by capillary electrophoresis on an ABI 3730 DNA Analyzer using standard conditions.

Sequences from each locus were assembled and aligned into contigs with Sequencher 4.6 (Gene Codes Corporation), and putative polymorphisms were then verified by eye on the chromatograms. Polymorphisms were considered for SNP assay development if both homozygote genotypes and the heterozygote genotype were observed at least once in the panel. This standard was adhered to in order to avoid identifying sequencing artifacts as polymorphisms, and to increase the likelihood that resulting SNP assays would have appropriate minor allele frequencies for future utilization. In cases where all variable sites were heterozygous, duplication of the gene in question was assumed owing to the partial diploidy of salmonids (salmonids diverged from a tetraploid ancestor 25-100 million years ago; Allendorf and Thorgaard 1984; Brieuc *et al.* 2014),

and it was no longer considered for development. In order to reduce the probability of SNP markers being in linkage disequilibrium, only one SNP per EST was selected for design. In consensus sequences where more than one candidate SNP occurred, the site with the greatest minor allele frequency in either the Scott Creek or Russian River populations was targeted. Consensus sequence information for polymorphic sites that met the above requirements, as well as TaqMan assay design requirements (e.g. more than 2 bp away from any other polymorphisms, more than 40 bp away from the end of the sequence) were then sent to Applied Biosystems for 5' exonuclease (TaqMan) assay design. Additionally, the species diagnostic assay was designed by aligning coho and Chinook salmon consensus sequences generated from three loci for which no polymorphisms were found in either species, and then searching for fixed allele differences within the alignment. The locus with the least number of fixed differences and that met the above requirements for TaqMan design was chosen for development. Finally, a BLAST search was performed on each consensus sequence to determine whether the EST matched, or was associated with, an identified gene, however this information was not used as a criteria for selecting the target variation (see Table 1.5).

In order to confirm marker quality and utility, each TaqMan assay was then validated by genotyping 470 coho salmon from five populations in California (Scott Creek; Lagunitas/Olema creeks; Noyo River; Russian River, Warm Springs Hatchery; and Klamath River, Iron Gate Hatchery), one population in Oregon (Nehalem River, Nehalem Fish Hatchery), three populations in Washington (Columbia River, Bonneville Hatchery; Quinault River, Quinault National Fish Hatchery; and Green River, Soos

Creek Hatchery), and one population in British Columbia (Fraser River, Chilliwack River Hatchery). SNP genotyping was performed using the Fluidigm EP1 Genotyping System. This high throughput approach uses 96.96 Dynamic Genotyping Arrays that employ nanofluidic circuitry to allow for the simultaneous analysis of 96 SNP assays by 96 DNA samples. Fluidigm SNP Genotyping analysis software was used to assemble genotype data, as well as to judge SNP plot quality. The program GENETIX 4.05 (Belkhir *et al.* 1996-2004) was used to estimate allele frequencies, expected (H_E) and observed (H_O) heterozygosity (Nei 1978) and the fixation index F_{ST} (Weir and Cockerham 1984). The program GENEPOP 4.0 (Rousset 2008) was used to calculate deviations from Hardy-Weinberg and linkage (gametic phase) equilibria.

1.3 Results

Of the 275 EST primer pairs chosen for SNP discovery, 248 produced PCR products that amplified as a single band for most individuals when electrophoresed in agarose. These ESTs were then sequenced, while those that did not amplify or that produced multiple bands were no longer considered. Of the 248 loci that went on to be sequenced, 234 produced sequence data at one or more individuals (Table 1.1). When taking into account both forward and reverse strands, an average of 33.7 (out of a maximum possible of 48) sequences per locus was attained. Most loci produced sequence data in the targeted length range with overlapping forward and reverse strands, however 56 loci generated a consensus sequence larger than 600 bp and 32 produced forward and

Table 1.1: Summary of EST sequencing effort to identify genetic variation in populations of Chinook salmon (*O. kisutch*) from the west coast of North America. The weighted estimates account for unobserved variation in consensus sequence derived from less than 24 individuals.

	Total	**Mean per locus [range]**
EST loci successfully sequenced	234	
Base-pairs sequenced (all fragments)	4,206,429	15813.64 [764-46246.5]
Length of consensus sequence (bp)	125,356	471.26 [72-1108]
Weighted consensus (bp)	118,199	444.36 [72-1108]
Number of observed substitutions	610	2.61 [0 - 12]
Number of SNPs (all three genotypes observed)	225	0.96 [0 - 6]
Loci with no variable sites	38	
Insertions/deletions (indels)	68	
Transitions (A-G or C-T)	279	
Transversions (A-C or G-C or A-T or G-T)	332	
Sites with 3 nucleotides observed	1	
Possible duplicated genes	6	
Total number of substitutions + indels	678	
Density of substitutions in consensus sequence	0.0049	
Density of substitutions in weighted consensus sequence	0.0052	

reverse sequences that did not overlap (consensus fragments for seven of these loci were also longer than targeted), indicating the presence of one or more intronic regions.

Approximately 4.2 MB of genomic sequence data was generated and aligned (mean 15.8 kb/locus) with a total consensus sequence length of approximately 125 kb (mean 471 bp/locus). Of the 234 EST loci that produced sequence data, 196 harbored one or more variable sites, comprising a total of 610 nucleotide substitutions or likely SNPs. The mean density of observed mutations in the ∼125 kb of total consensus sequence was calculated to be 0.0049, or about one substitution every 206 bp. In

order to correct for missing sequence data and thus more accurately estimate potential SNP density, the mean length of fragments making up the consensus sequence was weighted by the total number of individuals for which sequence was generated. The calculated weighted consensus sequence was approximately 118 KB (mean 444 bp/locus) and produced a mean substitution density of 0.0052, or about one substitution every 194 bp. The density of putative SNPs, or substitutions where all three genotypes were observed, in the consensus sequence was 0.0018, or about 1 SNP every 557 bp, while weighted the density was 0.0019, or about 1 SNP every 525 bp.

Mutations were only considered to be putative SNPs if all three genotypes were exhibited in the sequence data. Of the 225 putative SNPs uncovered in the sequencing effort, 117 met the TaqMan assay design criteria and were submitted for assay development. Of those, 111 were suitable for assay manufacture and were then used to genotype the validation populations. Of these 111 assays, 20 were discarded owing to their inability to reliably discriminate allelic variation. These assays produced plot results that either had no signal, had more than three clusters present, had only one apparent heterozygote cluster, had both homozygote clusters without a heterozygote cluster, or generally produced poorly discernible genotype clusters. A total of 91 SNP assays and the 1 species diagnostic assay then went on to be further validated and characterized. Assay primer/probe and variable base information, as well as GenBank dbGSS and NCBI dbSNP accession numbers, are listed in Table 1.2.

Table 1.2: Description of the 91 SNP assays and 1 coho/Chinook species diagnostic assay developed in this project with the target polymorphism, primer and probe sequences, length of the consensus sequence in base pairs (bp), and GenBank (dbGSS) and NCBI (dbSNP) accession numbers indicated.

Assay name	Targets	Primers (5′-3′)	Probes (5′-3′)	bp	dbGSS	dbSNP
Oki94903-192	T/A	F: CGTAAAAGTATTGCAGAAAATCCGAGTT R: GTACACAAAATGGATGACATTTGTAAGGT	VIC: TGCGAGAATATTGACTTTAAAT FAM: CGAGAATATTGACATTAAAT	318	KG772637	974293192
Oki95318-100	G/A	F: TCTGAATGATTATGCAGAACGTGTTCT R: TGCTGGAGATTATGAATTTGGTGTTGA	VIC: CCCTTTAACAGCCTTCA FAM: CCCTTTAACAACCTTCA	376	KG772638	974293193
Oki96127-66	T/G	F: CATCACAGTATAACAGGTAATGTTTCAGTTGAT R: GCAGCCTACTCCTAGTATGTGATGT	VIC: CTGCTATAAGTAATAATAATC FAM: TGCTATAAGTAAGAATAATC	514	KG772639	974293194
Oki96158-278	C/T	F: TTGTGACGGGATCCTTCATTGAG R: CACCTCGATCTCCTTTCGGTAAAA	VIC: ACGACCCCACAATAG FAM: AAATACGACCCTACAATAG	325	KG772640	974293195
Oki96222-70	G/C	F: GCTGCAGCCCAGGTAAGAAA R: GGATGAGAGAATTTAAAATTAGCTACAGTTGTAGT	VIC: ATGTTGCGACAATTAG FAM: TTGCGCCAATTAG	626	KG772641	974293196
Oki96376-63	A/T	F: TCTGCTACGTGTTTGACAGTGAAAA R: GTGTGGAACATGAACTCAAAACGTA	VIC: CACACTTTGAGAGTTTAAAAA FAM: ACACTTTGAGAGTTTAAAA	652	KG772642	974293197
Oki97660-149	A/C	F: CTGGTCCGGTCCAAATGATGTA R: GGGTAATCAGATTACACATAACTGACTACT	VIC: CCAACCCTGTATGTAACC FAM: CAACCCTGTCTGTAACC	444	KG772643	974293198
Oki97954-228	G/T	F: GACCAGAAGTTGCACAATAGTAAGGT R: TCGGCAAATACAGACGTGTTTCA	VIC: CACTGGTAAGCGTTGTTT FAM: ACACTGGTAAGCTTTGTTT	1006	KG772644	974293199
Oki100771-83	G/A	F: AGGTTTAAAGCCGTGTTCAATAGACA R: ACGTTTTGCAACAGAATTTTCGGAAT	VIC: CACCCCTGGCTCTGG FAM: CACCCCTGACTCTGG	401	KG772645	974293200
Oki100884-210	G/T	F: GCTGCGATTTAATGGTCATCCG R: TCTGGTCTTCCGCTGTGC	VIC: ACCCCTTAACATCGGGACA FAM: AACCCCTTAACATCGTGACA	449	KG772646	974293201
Oki100974-293	A/C	F: GGGATCCTGCAACACATTACAGT R: CTCTCACCAGAGAGATGTGGTTTT	VIC: ACAAAGTCAAATTTTTC FAM: CAAAGTCACATTTTTC	464	KG772647	974293202
Oki101119-1006	A/C	F: CAAAGACATAATCTGCAAATTACTAAAATGAGCAT R: GGTTCTGGACAAACCCATCCT	VIC: ATAGAACAATTATGGAATAAG FAM: AACAATTCTGGAATAAG	1108	KG772648	974293203
Oki101419-103	C/A	F: GGGTTCAATCCACTTCCTACTGTT R: TGACCCTCAATGACAAAATTACACATTTG	VIC: CTCCTACTTGCCTGTCTC FAM: CTCCTACTTGACTGTCTC	629	KG772649	974293204
Oki101554-359	A/G	F: CATCTGCTTCTCACGTTTTTAAAGGTAATTTATA R: CTCCATCACCACCACTACAATTGA	VIC: AAGTTTGCAAAAGTTG FAM: AAGTTTGCAGAAGTTG	410	KG772650	974293205
Oki101770-525	G/T	F: CTCTGCACTCTGTTATGCTACTGTA R: AGTTTTGCCTCTGTGCATCCT	VIC: AGAGAACACATGACTTAAT FAM: AGAGAACACATTACTTAAT	761	KG772651	974293206
Oki102213-604	G/A	F: CCACTTCACTCTGTGTCGTAACA R: GAAGTAATGTAGCAGATTAGGAGAATGAAGT	VIC: AGTAGTGAAAATGCTTTCC FAM: AGTAGTGAAAATACTTTCC	1052	KG772652	974293207
Oki102267-166	T/A	F: GACAGCCACGCAACAAGTTTATAT R: TCACTAAGTCGCTCTGGATAAAGC	VIC: TGCTAGCGGCTTCCACAT FAM: CTAGCGGCATCCACAT	439	KG772653	974293208
Oki102414-499	T/A	F: GGTCTCTTGTGGCGAAGTGTTAG R: CAATGGCAACCGGGATGAG	VIC: CCTGACACTGTGCTTC FAM: CTGACACAGTGCTTC	990	KG772654	974293209
Oki102457-67	G/T	F: GGCCAATGCTTGGAGTAAATACATC R: GGAGGCCCTACCAAGACATTTT	VIC: TGTAGAAATTACTTGAACCTT FAM: TTGTAGAAATTACTTTAACCTT	614	KG772655	974293210
Oki102801-511	C/T	F: TCACTGTCGGCTGCATAGG R: CTCTGTACGCAGTATCCTGACATTAC	VIC: ATGACGCCCCTGGGTG FAM: ATGACGCCCTTGGGTG	595	KG772656	974293211

Continued on next page

Table 1.2 – continued from previous page

Assay name	Targets	Primers (5′-3′)	Probes (5′-3′)	bp	dbGSS	dbSNP
Oki102867-667	G/T	F: GAGCCTCTATGGTGGCTATAATTATTTGT R: TCTTTGTTTTGCTTCTTAGGCACCTA	VIC: AATGAATGTTGAGACAGAAA FAM: AAATGAATGTTGATACAGAAA	732	KG772657	974293212
Oki103271-161	A/G	F: GAGACAACTAAGCAACCTCTGAGAA R: GGGTCATTTGAACCCTGATTTAAATCC	VIC: TCAGTTATGACAGATTTG FAM: AGTTATGGCAGATTTG	463	KG772658	974293213
Oki103577-70	T/A	F: CTGCAAATCACTCCTTCCATACTCT R: CTCCAGGTACTTAGCCCAATTTATTGTA	VIC: CCCTCCACATTTCACA FAM: CCTCCACAATTCACA	606	KG772659	974293214
Oki103713-182	A/G	F: GCTCTCCATCTTAGCCAGTCAATT R: GGGTGTACCCAGTCTTTATTTACAAT	VIC: AGTCCTGAAATGAAGATACA FAM: CCTGAAATGGAGATACA	242	KG772660	974293215
Oki104515-99	A/C	F: GGCTCCAGCCAGTCCATT R: GGAGACGTTGAAATTTGTGGAGAAAT	VIC: CAGCCCAAAGTCTG FAM: AGCCCCAAGTCTG	396	KG772661	974293216
Oki104519-45	G/A	F: TGAGAAAGAGGTGTGAGCGAAT R: ACGAGTCAGGCGCAAGTTA	VIC: CAGCGGCGTCATCA FAM: AGCGGCATCATCA	593	KG772662	974293217
Oki104569-261	A/C	F: CCGGTGATGTTGCCAATCTTG R: GGCAGTACGTCATCAAGCTCTATG	VIC: CGACGACATAACCTC FAM: CGACGACATCACCTC	581	KG772663	974293218
Oki105105-245	A/C	F: CGTATCAAGCATCAACGCCTTTT R: GCAAGGTTGGGCATGAGGAA	VIC: ATCTCTGTTTGAGAGTTACAT FAM: TCTGTTTGAGCGTTACAT	789	KG772664	974293219
Oki105115-49	A/C	F: GGTAACACTTTCGTAGCTTCATGGT R: CGTTGAGGTGACATGTAGCTATTCA	VIC: CGCTCATAGAAATAACTTAAA FAM: CGCTCATAGAAATACCTTAAA	402	KG772665	974293220
Oki105132-169	G/A	F: TGCCCTGAACCACATGTCAAA R: AAGAGTGGAATAAAAACCTGCAGACA	VIC: CCCTCAGTGCATCGTG FAM: CCCTCAGTACATCGTG	446	KG772666	974293221
Oki105235-460	A/C	F: CATAACTTTCAAACACAAACAGGTGCT R: TGTCTCATGTGCCCCAGTAAAATAA	VIC: AACTTTTGAACAATTTTT FAM: ACTTTTGAACCATTTTT	793	KG772667	974293222
Oki105385-521	C/A	F: CACACAGACTAACTTCTATCAGGCAAT R: CGTTCTCACATTTCAAACTGATATTAGTAGA	VIC: CACTTTTCTATGACCTTTTT FAM: ACTTTTCTATGAACTTTTT	565	KG772668	974293223
Oki105407-161	G/A	F: GGTTCAGGTTAGGAATGGTTGATGA R: GCCTATAGGATTATCTCTTGAAACAGCAA	VIC: AAATACAAACGAAGGCTG FAM: AAATACAAACAAAGGCTG	474	KG772669	974293224
Oki105897-298	A/T	F: TGCATCTACTGAATTCATTTCACAGTTT R: GCAACAGGGACACATTTCAGGAT	VIC: ACACAAATAATCCAATATGAAAT FAM: CACAAATAATCCAATTTGAAAT	387	KG772670	974293225
Oki106172-60	G/A	F: TGATAACTACTTGGCGTGTGTGT R: GCTGGCGGCAGTCAGT	VIC: TACGCACGATGGAGGGT FAM: CGCACGATAGAGGGT	465	KG772671	974293226
Oki106313-353	T/C	F: TCCCCATTACAATATGTCCCTGACA R: CCAAACCCTGTGTGAAAATGTGAAA	VIC: AGACACAACAGTTGGTAAGT FAM: CACAACAGTCGGTAAGT	699	KG772672	974293227
Oki106419-292	A/C	F: TCTGGATAAGAGCATCTGCTAAATGAC R: GATCCTGCCTTGTGCATACTTG	VIC: CAAGTTCCTTTCAGTTTTA FAM: CAAGTTCCTTTCCGTTTTA	768	KG772673	974293228
Oki106479-278	A/G	F: ACACTCAACACAGGCTCAATTAGTT R: TTGTTTACGTGCTGTAGATGTCTTCA	VIC: AAACACTACACAGATGCAT FAM: CACTACACGGATGCAT	626	KG772674	974293229
Oki107336-45	G/T	F: CGGTCCGAGAGGTCTTTGAG R: GGATGCTCAGAAGGGACTTGAA	VIC: TCAGAGCGGACTCCAG FAM: CTCAGAGCGTACTCCAG	472	KG772675	974293230
Oki107607-213	T/C	F: GCGTTGCACCTGATCTAGAGTAG R: CTTTGCAACCCTTGAGGAAAATGAA	VIC: CATGTTCTCCAGTGATGTA FAM: ATGTTCTCCAGCGATGTA	517	KG772676	974293231
Oki107974-46	A/G	F: GACCTCGAACTGAATCTTCTTTCCA R: CGCGCAAAACTGCCCAAA	VIC: CTGCGCCACTGGTG FAM: TGCGCCGCTGGTG	461	KG772677	974293232
Oki108505-331	G/A	F: GTCTGTGGCAGATGTATACTGAACA R: CTGATTTCCTTATATGAACTAACTGTAACTCAGT	VIC: TCATGTTGCGTTTATTT FAM: TCATGTTGCATTTATTT	554	KG772678	974293233

Continued on next page

Table 1.2 – continued from previous page

Assay name	Targets	Primers (5′-3′)	Probes (5′-3′)	bp	dbGSS	dbSNP
Oki109243-480	T/C	F: GTGGGATTTCCTGTTCTACCTGTAA R: AGGACGCAGTGTTGACATCATC	VIC: CGGCTCATTGCTGC FAM: CGGCTCACTGCTGC	505	KG772679	974293234
Oki109525-359	A/G	F: GATGATGATAGGCTCTTCCAGACAT R: CCACATGCTGCCCGCTTA	VIC: ATTCTCATTGGTAAGTTGTC FAM: TCATTGGTGAGTTGTC	1021	KG772680	974293235
Oki109651-152	T/A	F: GCTGTTGTCATATCATCCCGTTAAC R: CCTGATTTTGCCCACATTTCAAGAA	VIC: AGGTTAACATATGTAGCTTTAAAA FAM: AGGTTAACATATGTAGCTTAAAAA	641	KG772681	974293236
Oki109874-122	G/T	F: GAATGGGTAGCTAAACATGAATAGGCA R: TCAGTTTTGAACAGCATTGATTTTGTCA	VIC: CCTGACACTAACGTCTAAT FAM: CCTGACACTAACTTCTAAT	382	KG772682	974293237
Oki109894-418	C/A	F: GAGGTACCAGGTGAGCTCAAAC R: GACCAGACCATAGACATTGCCTTA	VIC: TGACTGGTTGCTTGAACA FAM: TGACTGGTTGATTGAACA	538	KG772683	974293238
Oki110064-418	T/C	F: CGTTCATTACGTAGCCTTTTAGGGATT R: GCTTTGCACTCTTAACATAATAGTTTTGGA	VIC: TGCAAGGGACTTAGCT FAM: TGCAAGGGACCTAGCT	644	KG772684	974293239
Oki110078-191	T/G	F: ACACACTCGCTCTAGCTCCTA R: TGTTCTGAAAGGGCCACTTGAATA	VIC: TGACTTTTACACTCTAATGTC FAM: TTTTACACTCGAATGTC	474	KG772685	974293240
Oki110381-77	T/C	F: CAACTTCCCCAAAATCAAACAGTGA R: CATCGAGGGTCGCCTATCTG	VIC: CCGTATCTTAAACCC FAM: CGTATCCTAAACCC	193	KG772686	974293241
Oki110689-43	G/T	F: AAATCTCCTATGTGGTGTGCC R: CACTCCTACAGAACAGATTCAAGGT	VIC: TGGGTCATTTTCTATGGCATC FAM: TGGGTCATTTTCTATTGCATC	376	KG772687	974293242
Oki111681-407	T/A	F: CGGTCATGTTCATCCCATTGGAA R: TGCATGATGCATAGCACCTAATATGAA	VIC: CTGTGCGCTTTTGCAG FAM: TGTGCGCTATTGCAG	482	KG772688	974293243
Oki113457-324	A/G	F: GTGGTCCAGGCTGACAGT R: AGCACCAGCAACTCTTTCCT	VIC: CCCTCTGTGACTCGTT FAM: CCTCTGTGGCTCGTT	671	KG772689	974293244
Oki113979-170	C/T	F: CCCACACTCCCAAAGTTACATATCT R: TGGCCATCGATGTGGGATTTT	VIC: CCTTCAGACAACGGAAAA FAM: CCTTCAGACAATGGAAAA	246	KG772690	974293245
Oki114315-360	T/C	F: CTGAGGAGGGAGGCAGTAGT R: GACCCTTGACCTGTGACATCA	VIC: CACCATATGCTATTACTC FAM: CCATATGCCATTACTC	534	KG772691	974293246
Oki114448-101	G/A	F: AAATCCTTTAGAGTACAGAAAATGTTCGATCA R: TCCTGCTGGACTAGGCTAACA	VIC: AAGCTGACGGGTGCAGT FAM: AGCTGACGGATGCAGT	528	KG772692	974293247
Oki114587-309	A/C	F: TGTCACAAATGATCTGCAAAACACATT R: GCAGTATCCAGGTCAAGAAAACCA	VIC: CTGTCTCAAAAATGAATAGTA FAM: CTGTCTCAAAAATGCATAGTA	485	KG772693	974293248
Oki115987-366	G/A	F: GTCGTTATGTGGGCTTTCTAATAGC R: CCGTAGTACAGTGTATCGAGTGGAA	VIC: TATGGCATTCGGTGAACC FAM: TGGCATTCAGTGAACC	591	KG772694	974293249
Oki116362-411	A/T	F: GAGGCTGTTCTTCTGGATGCA R: CCCACCTTGCCTGCTATAGTT	VIC: ATTAACATTACACAAAACATT FAM: AATTAACATTACACATAACATT	641	KG772695	974293250
Oki116865-244	A/G	F: AGCTGATAGATAAGAAAAATCTTCATAAAACCCATT R: ACCGGGTAAAAGTAGGTCCATAGAA	VIC: ATCCTTTATGTTATAGATTTT FAM: CCTTTATGTTATGGATTTT	489	KG772696	974293251
Oki117043-374	G/A	F: TGTTTGCAGTCTTCCTCTTAACCA R: TGCAAGGCCCACTCATCTATG	VIC: CAATTGTTGTTATCGGTGTCC FAM: CAATTGTTGTTATCAGTGTCC	614	KG772697	974293252
Oki117144-64	C/T	F: AACCACATGCTCCACAGACA R: GCCTGGTGGCCTTGCT	VIC: TCCTCCACCACTGAGC FAM: CTCCTCCACTACTGAGC	465	KG772698	974293253
Oki117286-291	G/T	F: TCTCAAGTCAGAATAATGCCAAAAGTGA R: TGATCAGGTAATGCTGCTTGAGAATATTTTA	VIC: CAGGTATGGTAGAGGTC FAM: CAGGTATGTTAGAGGTC	456	KG772699	974293254
Oki117742-259	G/A	F: GGTGGAGGAAAAGGAGACTTCTG R: CAGAACTCCACTGAGACGTTGTT	VIC: CTTGTCCACGTGCAGAC FAM: CTTGTCCACATGCAGAC	407	KG772700	974293255

Continued on next page

Table 1.2 – continued from previous page

Assay name	Targets	Primers (5′-3′)	Probes (5′-3′)	bp	dbGSS	dbSNP
Oki117815-369	T/G	F: CATTTGCTCTGTCTCCCACTCT R: CAGCAAGCCTCACAGAAACC	VIC: ATGTTTGGCTTTTTTTGTTG FAM: TTTGGCTTTGTTTGTTG	393	KG772701	974293256
Oki118152-314	A/G	F: CTTCGCTGTCTACGGTACAAACA R: TCCAGAACCTGTTGTTGGGAAAT	VIC: CAACTTCTTTAGCTATCTACTC FAM: ACTTCTTTAGCTGTCTACTC	710	KG772702	974293257
Oki118175-264	T/A	F: CCCAAATCCTAACCATTTTGTGGATGA R: GTAGGTAGAACCTTACCCCTTACCA	VIC: CTATAGGCTCAGTCATTTT FAM: TAGGCTCAGACATTTT	411	KG772703	974293258
Oki118654-330	G/A	F: CCCTGCGCATCCACAAC R: AGCTTGCTGACCTGGTTGAG	VIC: CCTGGTCGGAGGTGT FAM: CCTGGTCAGAGGTGT	454	KG772704	974293259
Oki119108-320	G/A	F: CTGCGGGAGCTACACTCA R: TGGCAAGGACATGTGTGACT	VIC: CCTGGACGCGGACTG FAM: CCTGGACGCAGACTG	830	KG772705	974293260
Oki120024-226	A/C	F: AAACGAGGCCTTAAATATCTCAACCA R: GAACGGCACGCTCTAAAGTG	VIC: TGGTGTGAAACTATAATCA FAM: TGGTGTGAAACTCTAATCA	296	KG772706	974293261
Oki120255-113-sppID	T/C	F: CAGGCTACAGGGACTTTACAATGG R: ATTGAAGGGTGGAATTGAAGTTAGCT	VIC: ATGAAGCTACCCTAGTAACT FAM: AAGCTACCCCAGTAACT	113	KG772707	974293262
Oki121006-412	G/T	F: GGCAAATCTGTTGTGATCAAATGCT R: CAAAGTTTAAAAAGCAACAATTGCAGTACTC	VIC: AAGGGTTGAGGTTTTT FAM: AAGGGTTGATGTTTTT	498	KG772708	974293263
Oki122593-430	A/G	F: CAAATTCTGGGTAAAGTGTGCCTTT R: AGCGCTGAACATGCACAATAAATG	VIC: TGTTTCTCACCAGTCTTTT FAM: TTCTCACCGGTCTTTT	747	KG772709	974293264
Oki123044-68	G/T	F: TCATTTCACTGCCATATCAGTATTGGG R: GCGTGGCAAAAAAGTAAGAATGCTA	VIC: CGGAATCACTTTAAAAACGTTTT FAM: AATCACTTTAAAAACTTTTT	723	KG772710	974293265
Oki123205-88	T/C	F: GCAACAAAGACAACTTTAAACCAAAAACTT R: CACCTGCCTGGGAAAATTATGG	VIC: CCTACTCCATTGCTTTT FAM: CTACTCCACTGCTTTT	311	KG772711	974293266
Oki123470-92	A/G	F: CCTTGGTTTATTGTGTTCTTGCGTAT R: ACCTGTGTGGGAGATGTGTGTATAT	VIC: TCTAACCAGAACAAGACC FAM: TAACCAGAGCAAGACC	450	KG772712	974293267
Oki123921-90	T/C	F: GCGACGGAGATGAAGGTATGAATTA R: CCTGAACCTTAGGGAACATCACATC	VIC: AGTTGCAAGATAATTTACA FAM: AGTTGCAAGATAACTTACA	912	KG772713	974293268
Oki124162-62	A/T	F: CCAGGTGTCAGGTGATAAGGCTATA R: ACCTTTTTGTGCACTTTACATGTTGTC	VIC: CACTCAGCAAAAATATAAA FAM: CACTCAGCAAAATTATAAA	190	KG772714	974293269
Oki125998-340	A/C	F: CAGGTCTTCTTCTAAGCAGTCTCAT R: CCCAATAAAGGAACACATCGACATC	VIC: CAGGCCCTCAGTATGT FAM: CAGGCCCTCCGTATGT	467	KG772715	974293270
Oki126160-142	CT/TG	F: TCTACACAATTATTTTGCTCACCAACAC R: TCCCTTGTACAAAATGGCTGCTAAA	VIC: TGTTTGATCCTAAATTGC FAM: TTTGATCTGAAATTGC	649	KG772716	974293271
Oki127236-383	A/C	F: TTGTCTCCCAGGATATTGCTGTTTT R: GCTGGAAATATCACATTCCTCAAAATGTAG	VIC: CAGACAAACAAACAAAAA FAM: CAGACAAACAAACCAAAA	565	KG772717	974293272
Oki127760-301	T/C	F: ACAAATCGAGTACTGCAAAAAAATATACAAACAA R: AGGACTCGAACATTCCGTTTCTG	VIC: TTGCATTTCGATCAAAAG FAM: TGCATTTCGACCAAAAG	677	KG772718	974293273
Oki128302-547	T/C	F: AGAGCACAGAAGATACTTAGAAAAGAAAAACA R: GGAGTGGGATAAGAGGGTTTACAGT	VIC: AACATTCATCCCTATTTGTT FAM: ATTCATCCCCATTTGTT	961	KG772719	974293274
Oki128693-70	G/T	F: GGGTCAAGAGGGAAAGGC R: CTAAGAATTCCATGTCCCCACCAA	VIC: CTGGAAAATGTAATTCT FAM: ACTGGAAAATTTAATTCT	513	KG772720	974293275
Oki128757-232	A/T	F: ATGTTTTCAAAACGTTATTGTACGATGAC R: AGACCGGGTGCAACACAA	VIC: TGATTTGGGAAAGCTTT FAM: TGATTTGGGATAGCTTT	375	KG772721	974293276
Oki128851-185	A/G	F: TCTCCAGTCTCCCAGTTTACACA R: CCCTCTCCTGGAGAAAACAGAATAA	VIC: ATGGATAACCAGACTCCCT FAM: TGGATAACCAGGCTCCCT	329	KG772722	974293277

Continued on next page

Table 1.2 – continued from previous page

Assay name	Targets	Primers (5′-3′)	Probes (5′-3′)	bp	dbGSS	dbSNP
Oki129870-552	A/G	F: GATACTCAATTGATGTGGTCCCCAAT R: AAAAAGGAACAGAAATGTTGCGTCAT	VIC: CATCATGCTATTATGTTTTT FAM: CATGCTATTGTGTTTTT	923	KG772723	974293278
Oki130295-48	A/G	F: CAAACAATGTTACATGAGCACAAATATAAATG R: GCACAAGACATACAGTATGAACAAGAAA	VIC: CTGTTAAGCAATAAACAGTT FAM: TGTTAAGCAATGAACAGTT	440	KG772724	974293279
Oki130524-184	G/A	F: CCAGGCGCTGCCATTAC R: GTCTGCTCTGCTGTGTGCTTA	VIC: CATGAGTCAGCTTCCT FAM: CATGAGTCAACTTCCT	378	KG772725	974293280
Oki131147-353	A/C	F: TGTCAAAAGTAGTGCACGAAATAAGGA R: CACCTGTGTCTGTCTGTGTTTCTTA	VIC: TGCGTCCCAAAAAT FAM: TTGCGTCCCAAACAT	436	KG772726	974293281
Oki131460-243	G/T	F: AGGAACTTGGAACAATCGAGCAA R: GCTTGGCCATGATAACCTCAAAAT	VIC: CAGGGATGTAGAAGGAA FAM: CAGGGATGTATAAGGAA	607	KG772727	974293282
Oki131802-368	A/G	F: TGTTTATCACCTGATTGTCTCATGGC R: TGGATTGCTTACTTACTCATCTTCATTCTC	VIC: CAATTGTCACAAGAAGCT FAM: TTGTCACGAGAAGCT	493	KG772728	974293283

All 92 assays were used to genotype 470 coho salmon individuals sampled from ten populations throughout California, Oregon, Washington, and British Columbia, in order to assess assay genetic variability and efficacy throughout different regions of the species' North American range. A summary of the population genetic variability of the 91 validated SNP loci (the species diagnostic assay is not included in this table because it is characteristically monomorphic for one allele in coho salmon and for the alternative allele in Chinook salmon) can be found in Table 1.3 and Table 1.4. Of the 470 samples genotyped, 39 were removed from further analyses due to missing data at 20 or more loci, leaving a total 431 individuals to be analyzed. Minor allele frequency for individual loci ranged from 0.011 to 0.500, while mean minor allele frequency varied from 0.266 in Scott Creek to 0.398 in Quinault River, and averaged 0.349 over all populations. The proportion of polymorphic loci was 91.3% in all populations and ranged from 87.0% in the Nehalem River Hatchery population to 94.4% in the Lagunitas/Olema creeks population. Twenty-two of the markers were fixed in one or more of the population test groups, however all markers displayed variability in at least one of the ten populations. For each variable locus, expected heterozygosity (H_E) ranged from 0.021-0.507 (mean=0.319) and observed heterozygosity (H_O) ranged from 0.022-0.706 (mean=0.315). Mean H_E varied from 0.284 (Fraser River) to 0.344 (Noyo River), while mean H_O ranged from 0.271 (Fraser River) to 0.354 (Scott Creek). Mean F_{ST} for all loci in the California populations was 0.104, and ranged from 0.003 to 0.405 at individual loci over the five populations. In the Oregon, Washington and British Columbia populations, Mean F_{ST} for all loci was 0.073, and ranged from 0 to 0.438. Over all 10

populations, mean F_{ST} for all loci was 0.131, and values varied from 0.006 to 0.445 (data not shown). In total, five loci were not in Hardy-Weinberg equilibrium in all populations, three of which (Oki103577-70, Oki119108-320 and Oki131460-243) deviated from equilibrium in one population and two of which (Oki109525-359 and Oki129870-552) deviated from equilibrium in two populations (Table 1.3 and Table 1.4). Estimates of linkage disequilibrium between locus pairs varied when the Russian River test sample was included in the analysis versus when it was excluded. When included, 88 locus pairs were in complete disequilibrium, 63 of which were in significant linkage disequilibrium even though each locus was designed from a different gene. This is higher than expected by chance alone with a significance value of ≤ 0.001. When the Russian River test sample was excluded, only three locus pairs were in significant linkage disequilibrium based on a significance value of 0.001 or less. This phenomenon is most likely the result of a small effective population size consisting of closely related individuals due to near extirpation from the Russian River watershed in the last few decades (Garza and Gilbert-Horvath 2003, Bucklin *et al.* 2007).

Table 1.3: Summary statistics for 91 SNP loci in five coho salmon populations within California. N is the number of individuals genotyped. H_E is expected (unbiased) heterozygosity and H_O is observed heterozygosity. F_{ST} is over all five populations. AF is the observed frequency of the minor allele from the Scott Creek individuals in each population. Asterisks (*) indicate significant ($p < 0.001$) deviations from Hardy-Weinberg equilibrium. Populations are arranged from south to north. WSH = Warm Springs Hatchery; IGH = Iron Gate Hatchery.

	Scott Ck.			**Lag./Olema**			**Russian R. (WSH)**			**Noyo R.**			**Klamath R. (IGH)**			
	N=44			**N=45**			**N=47**			**N=44**			**N=44**			
Assay	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	F_{ST}
Oki94903-192	0.352	0.46	0.43	0.478	0.50	0.51	0.500	0.51	0.28	0.386	0.48	0.55	0.432	0.50	0.41	0.003
Oki95318-100	0.193	0.32	0.30	0.144	0.25	0.24	0.053	0.10	0.11	0.284	0.41	0.52	0.337	0.45	0.44	0.067
Oki96127-66	0.209	0.33	0.37	0.267	0.40	0.31	0.170	0.29	0.30	0.114	0.20	0.23	0.034	0.07	0.02	0.044
Oki96158-278	0.443	0.50	0.52	0.378	0.48	0.49	0.304	0.43	0.43	0.295	0.42	0.50	0.761	0.37	0.39	0.145
Oki96222-70	0.409	0.49	0.41	0.389	0.48	0.69	0.521	0.50	0.49	0.375	0.47	0.48	0.256	0.39	0.41	0.027
Oki96376-63	0.432	0.50	0.55	0.500	0.51	0.51	0.098	0.18	0.15	0.318	0.44	0.55	0.205	0.33	0.27	0.114
Oki97660-149	0.000	0.00	0.00	0.033	0.07	0.07	0.277	0.40	0.47	0.011	0.02	0.02	0.136	0.24	0.18	0.173
Oki97954-228	0.273	0.40	0.36	0.400	0.49	0.49	0.383	0.48	0.60	0.318	0.44	0.45	0.239	0.37	0.25	0.011
Oki100771-83	0.050	0.10	0.10	0.000	0.00	0.00	0.000	0.00	0.00	0.000	0.00	0.00	0.000	0.00	0.00	0.040
Oki100884-210	0.314	0.44	0.44	0.022	0.04	0.04	0.170	0.29	0.21	0.034	0.07	0.07	0.136	0.24	0.27	0.110
Oki100974-293	0.205	0.33	0.32	0.364	0.47	0.50	0.830	0.29	0.21	0.534	0.50	0.57	1.000	0.00	0.00	0.405
Oki101119-1006	0.114	0.20	0.23	0.122	0.22	0.24	0.000	0.00	0.00	0.045	0.09	0.09	0.183	0.30	0.32	0.044
Oki101419-103	0.430	0.50	0.53	0.511	0.51	0.52	0.609	0.48	0.30	0.535	0.50	0.51	0.693	0.43	0.30	0.028
Oki101554-359	0.330	0.45	0.52	0.456	0.50	0.38	0.128	0.23	0.21	0.261	0.39	0.25	0.000	0.00	0.00	0.152
Oki101770-525	0.023	0.04	0.05	0.078	0.15	0.16	0.000	0.00	0.00	0.000	0.00	0.00	0.035	0.07	0.07	0.033
Oki102213-604	0.345	0.46	0.55	0.678	0.44	0.42	0.649	0.46	0.36	0.600	0.49	0.45	0.951	0.09	0.10	0.182
Oki102267-166	0.214	0.34	0.33	0.378	0.48	0.49	0.330	0.45	0.53	0.261	0.39	0.43	0.279	0.41	0.42	0.009
Oki102414-499	0.239	0.37	0.30	0.318	0.44	0.41	0.553	0.50	0.43	0.364	0.47	0.50	0.244	0.37	0.34	0.069
Oki102457-67	0.284	0.41	0.39	0.500	0.51	0.47	0.755	0.37	0.36	0.670	0.45	0.57	0.857	0.25	0.19	0.203
Oki102801-511	0.163	0.28	0.18	0.289	0.42	0.44	0.340	0.45	0.43	0.159	0.27	0.27	0.000	0.00	0.00	0.095

Continued on next page

Table 1.3 – continued from previous page

	Scott Ck.			**Lag./Olema**			**Russian R. (WSH)**			**Noyo R.**			**Klamath R. (IGH)**			
	N=44			**N=45**			**N=47**			**N=44**			**N=44**			
Assay	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	F_{ST}
Oki102867-667	0.398	0.48	0.48	0.182	0.30	0.36	0.446	0.50	0.63	0.267	0.40	0.44	0.000	0.00	0.00	0.145
Oki103271-161	0.398	0.48	0.52	0.398	0.48	0.61	0.553	0.50	0.55	0.352	0.46	0.52	0.580	0.49	0.52	0.033
Oki103577-70	0.455	0.50	0.45	0.709	0.42	0.44	0.815	0.30	0.28	0.625	0.47	0.43	0.291	0.42	0.21*	0.167
Oki103713-182	0.116	0.21	0.23	0.111	0.20	0.22	0.106	0.19	0.21	0.163	0.28	0.28	0.619	0.48	0.52	0.350
Oki104515-99	0.341	0.45	0.50	0.205	0.33	0.27	0.479	0.50	0.36	0.261	0.39	0.39	0.705	0.42	0.36	0.157
Oki104519-45	0.432	0.50	0.59	0.511	0.51	0.49	0.372	0.47	0.40	0.477	0.50	0.55	0.663	0.45	0.53	0.034
Oki104569-261	0.068	0.13	0.14	0.011	0.02	0.02	0.000	0.00	0.00	0.068	0.13	0.14	0.182	0.30	0.23	0.084
Oki105105-245	0.057	0.11	0.11	0.011	0.02	0.02	0.074	0.14	0.11	0.091	0.17	0.18	0.295	0.42	0.41	0.151
Oki105115-49	0.500	0.51	0.50	0.311	0.43	0.49	0.372	0.47	0.40	0.557	0.50	0.61	0.545	0.50	0.32	0.037
Oki105132-169	0.386	0.48	0.45	0.189	0.31	0.33	0.564	0.50	0.57	0.443	0.50	0.48	0.671	0.45	0.51	0.122
Oki105235-460	0.286	0.41	0.57	0.250	0.38	0.45	0.566	0.50	0.45	0.529	0.51	0.29	0.446	0.50	0.46	0.071
Oki105385-521	0.318	0.44	0.27	0.476	0.50	0.38	0.522	0.50	0.51	0.330	0.45	0.39	0.183	0.30	0.22	0.064
Oki105407-161	0.330	0.45	0.57	0.284	0.41	0.25	0.723	0.40	0.21	0.443	0.50	0.52	0.511	0.51	0.52	0.116
Oki105897-298	0.432	0.50	0.59	0.733	0.40	0.40	0.713	0.41	0.36	0.833	0.28	0.24	0.605	0.48	0.32	0.098
Oki106172-60	0.151	0.26	0.26	0.411	0.49	0.56	0.606	0.48	0.49	0.477	0.50	0.59	0.545	0.50	0.41	0.112
Oki106313-353	0.419	0.49	0.60	0.489	0.51	0.40	0.660	0.45	0.47	0.511	0.51	0.52	0.648	0.46	0.57	0.034
Oki106419-292	0.372	0.47	0.51	0.478	0.50	0.56	0.351	0.46	0.36	0.352	0.46	0.43	0.216	0.34	0.25	0.026
Oki106479-278	0.318	0.11	0.11	0.432	0.36	0.42	0.511	0.24	0.23	0.432	0.41	0.33	0.955	0.39	0.47	0.040
Oki107336-45	0.307	0.43	0.48	0.278	0.41	0.42	0.574	0.49	0.38	0.545	0.50	0.50	0.477	0.50	0.72	0.066
Oki107607-213	0.136	0.24	0.27	0.044	0.09	0.09	0.255	0.38	0.26	0.068	0.13	0.14	0.318	0.44	0.32	0.087
Oki107974-46	0.398	0.48	0.48	0.333	0.45	0.57	0.734	0.39	0.28	0.523	0.50	0.58	0.477	0.50	0.49	0.089
Oki108505-331	0.443	0.50	0.52	0.500	0.51	0.50	0.628	0.47	0.40	0.407	0.49	0.53	0.837	0.28	0.23	0.107
Oki109243-480	0.352	0.46	0.66	0.067	0.13	0.13	0.467	0.50	0.28	0.233	0.36	0.37	0.105	0.19	0.21	0.143
Oki109525-359	0.303	0.43	0.45	0.310	0.43	0.29	0.476	0.50	0.27	0.434	0.50	0.29	0.333	0.45	0.51	0.012
Oki109651-152	0.337	0.45	0.40	0.656	0.46	0.51	0.826	0.29	0.26	0.636	0.47	0.41	0.500	0.51	0.64	0.129

Continued on next page

Table 1.3 – continued from previous page

	Scott Ck.			**Lag./Olema**			**Russian R. (WSH)**			**Noyo R.**			**Klamath R. (IGH)**			
	N=44			**N=45**			**N=47**			**N=44**			**N=44**			
Assay	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	F_{ST}
Oki109874-122	0.298	0.42	0.55	0.159	0.27	0.27	0.000	0.00	0.00	0.307	0.43	0.43	0.057	0.11	0.11	0.120
Oki109894-418	0.477	0.50	0.45	0.656	0.46	0.33	0.660	0.45	0.55	0.648	0.46	0.34	0.693	0.43	0.52	0.020
Oki110064-418	0.341	0.45	0.45	0.067	0.13	0.13	0.085	0.16	0.09	0.136	0.24	0.27	0.174	0.29	0.26	0.090
Oki110078-191	0.500	0.51	0.45	0.533	0.50	0.58	0.319	0.44	0.43	0.337	0.45	0.53	0.298	0.42	0.40	0.041
Oki110381-77	0.440	0.50	0.60	0.557	0.50	0.66	0.468	0.50	0.51	0.511	0.51	0.52	0.966	0.07	0.07	0.176
Oki110689-43	0.057	0.11	0.11	0.011	0.02	0.02	0.085	0.16	0.17	0.012	0.02	0.02	0.314	0.44	0.44	0.227
Oki111681-407	0.091	0.17	0.18	0.022	0.04	0.04	0.064	0.12	0.09	0.023	0.04	0.05	0.384	0.48	0.44	0.291
Oki113457-324	0.463	0.50	0.59	0.595	0.49	0.57	0.745	0.38	0.43	0.659	0.45	0.45	0.651	0.46	0.42	0.036
Oki113979-170	0.057	0.11	0.11	0.011	0.02	0.02	0.064	0.12	0.09	0.000	0.00	0.00	0.058	0.11	0.07	0.006
Oki114315-360	0.455	0.50	0.55	0.733	0.40	0.36	0.723	0.40	0.38	0.558	0.50	0.47	0.920	0.15	0.16	0.131
Oki114448-101	0.307	0.43	0.48	0.633	0.47	0.42	0.739	0.39	0.39	0.593	0.49	0.35	0.640	0.47	0.44	0.105
Oki114587-309	0.395	0.48	0.47	0.568	0.50	0.45	0.713	0.41	0.40	0.465	0.50	0.60	0.170	0.29	0.25	0.159
Oki115987-366	0.000	0.00	0.00	0.011	0.02	0.02	0.074	0.14	0.15	0.011	0.02	0.02	0.182	0.30	0.36	0.119
Oki116362-411	0.273	0.40	0.41	0.633	0.47	0.51	0.128	0.23	0.21	0.477	0.50	0.45	0.464	0.50	0.45	0.156
Oki116865-244	0.291	0.42	0.44	0.311	0.43	0.36	0.660	0.45	0.38	0.463	0.50	0.39	0.615	0.48	0.56	0.106
Oki117043-374	0.023	0.04	0.05	0.022	0.04	0.04	0.202	0.33	0.32	0.000	0.00	0.00	0.011	0.02	0.02	0.216
Oki117144-64	0.489	0.51	0.43	0.356	0.46	0.44	0.372	0.47	0.45	0.349	0.46	0.51	0.773	0.36	0.36	0.121
Oki117286-291	0.205	0.33	0.23	0.400	0.49	0.44	0.011	0.02	0.02	0.114	0.20	0.18	0.170	0.29	0.25	0.140
Oki117742-259	0.136	0.24	0.27	0.033	0.07	0.07	0.000	0.00	0.00	0.057	0.11	0.11	0.286	0.41	0.33	0.147
Oki117815-369	0.136	0.24	0.18	0.144	0.25	0.29	0.415	0.49	0.32	0.372	0.47	0.37	0.080	0.15	0.11	0.119
Oki118152-314	0.366	0.47	0.63	0.489	0.51	0.48	0.819	0.30	0.32	0.636	0.47	0.41	0.640	0.47	0.40	0.112
Oki118175-264	0.330	0.45	0.39	0.352	0.46	0.52	0.213	0.34	0.34	0.558	0.50	0.51	0.023	0.05	0.05	0.165
Oki118654-330	0.409	0.49	0.64	0.568	0.50	0.45	0.447	0.50	0.47	0.636	0.47	0.50	0.465	0.50	0.47	0.024
Oki119108-320	0.057	0.11	0.11	0.000	0.00	0.00	0.064	0.12	0.00*	0.000	0.00	0.00	0.012	0.02	0.02	0.015
Oki120024-226	0.443	0.50	0.34	0.633	0.47	0.56	0.574	0.49	0.30	0.602	0.48	0.52	0.693	0.43	0.43	0.023

Continued on next page

Table 1.3 – continued from previous page

	Scott Ck.			**Lag./Olema**			**Russian R. (WSH)**			**Noyo R.**			**Klamath R. (IGH)**			
	N=44			**N=45**			**N=47**			**N=44**			**N=44**			
Assay	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	F_{ST}
Oki121006-412	0.034	0.07	0.07	0.078	0.15	0.16	0.000	0.00	0.00	0.057	0.11	0.11	0.023	0.04	0.05	0.013
Oki122593-430	0.405	0.49	0.48	0.544	0.50	0.47	0.404	0.49	0.55	0.307	0.43	0.52	0.163	0.28	0.28	0.074
Oki123044-68	0.443	0.50	0.66	0.389	0.48	0.42	0.576	0.49	0.37	0.321	0.44	0.60	0.288	0.41	0.43	0.045
Oki123205-88	0.000	0.00	0.00	0.067	0.13	0.13	0.120	0.21	0.24	0.057	0.11	0.11	0.057	0.11	0.11	0.021
Oki123470-92	0.432	0.50	0.55	0.651	0.46	0.42	0.826	0.29	0.30	0.893	0.19	0.21	1.000	0.00	0.00	0.266
Oki123921-90	0.034	0.07	0.07	0.078	0.15	0.11	0.181	0.30	0.28	0.136	0.24	0.27	0.023	0.04	0.05	0.044
Oki124162-62	0.102	0.19	0.16	0.433	0.50	0.51	0.191	0.31	0.26	0.214	0.34	0.29	0.727	0.40	0.27	0.287
Oki125998-340	0.466	0.50	0.48	0.611	0.48	0.51	0.330	0.45	0.45	0.432	0.50	0.45	0.068	0.13	0.14	0.155
Oki126160-142	0.500	0.51	0.68	0.144	0.25	0.29	0.340	0.45	0.38	0.216	0.34	0.34	0.205	0.33	0.32	0.093
Oki127236-383	0.000	0.00	0.00	0.170	0.29	0.20	0.117	0.21	0.11	0.093	0.17	0.19	0.091	0.17	0.18	0.025
Oki127760-301	0.000	0.00	0.00	0.000	0.00	0.00	0.021	0.04	0.04	0.045	0.09	0.09	0.174	0.29	0.35	0.147
Oki128302-547	0.125	0.22	0.16	0.133	0.23	0.18	0.500	0.51	0.45	0.159	0.27	0.27	0.453	0.50	0.49	0.163
Oki128693-70	0.286	0.41	0.43	0.375	0.47	0.34	0.255	0.38	0.38	0.349	0.46	0.51	0.105	0.19	0.21	0.041
Oki128757-232	0.068	0.13	0.14	0.189	0.31	0.38	0.415	0.49	0.36	0.000	0.00	0.00	0.182	0.30	0.32	0.180
Oki128851-185	0.398	0.48	0.57	0.489	0.51	0.49	0.435	0.50	0.43	0.352	0.46	0.43	0.326	0.44	0.37	0.006
Oki129870-552	0.057	0.11	0.11	0.144	0.25	0.07*	0.021	0.04	0.04	0.159	0.27	0.27	0.125	0.22	0.20	0.022
Oki130295-48	0.034	0.07	0.07	0.311	0.43	0.40	0.457	0.50	0.49	0.273	0.40	0.27	0.570	0.50	0.40	0.163
Oki130524-184	0.221	0.35	0.35	0.170	0.29	0.20	0.426	0.49	0.51	0.452	0.50	0.48	0.488	0.51	0.56	0.077
Oki131147-353	0.352	0.46	0.39	0.686	0.44	0.30	0.596	0.49	0.55	0.726	0.40	0.36	0.779	0.35	0.35	0.115
Oki131460-243	0.102	0.19	0.20	0.189	0.31	0.38	0.617	0.48	0.26*	0.186	0.31	0.37	0.581	0.49	0.47	0.250
Oki131802-368	0.000	0.00	0.00	0.033	0.07	0.07	0.000	0.00	0.00	0.023	0.04	0.05	0.081	0.15	0.16	0.037
Mean	0.266	0.34	0.35	0.319	0.34	0.34	0.380	0.34	0.31	0.325	0.34	0.35	0.382	0.32	0.30	0.105
Polymorphic Loci (%)		90.7			94.4			91.7			93.5			93.5		

Table 1.4: Summary statistics for 91 SNP loci in five coho salmon populations ranging from Oregon to British Columbia. N is the number of individuals genotyped. H_E is expected (unbiased) heterozygosity and H_O is observed heterozygosity. F_{ST} is over all five populations. AF is the observed frequency of the minor allele from the Scott Creek individuals (Table 1.3) in each population. Asterisks (*) indicate significant ($p < 0.001$) deviations from Hardy-Weinberg equilibrium. Populations are arranged south to north. NFH = Nehalem Fish Hatchery; BH = Bonneville Hatchery; QNFH = Quinault National Fish Hatchery; SCH = Soos Creek Hatchery; CRH = Chilliwack River Hatchery.

	Nehalem R. (NFH) N=44			**Columbia R. (BH) N=38**			**Quinault R. (QNFH) N=45**			**Green R. (SCH) N=39**			**Fraser R. (CRH) N=41**			
Assay	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	F_{ST}
Oki94903-192	0.750	0.38	0.36	0.864	0.24	0.21	0.654	0.46	0.54	0.500	0.51	0.47	0.476	0.50	0.51	0.096
Oki95318-100	0.640	0.47	0.44	0.681	0.44	0.53	0.756	0.37	0.44	0.628	0.47	0.44	0.524	0.50	0.32	0.021
Oki96127-66	0.489	0.51	0.57	0.487	0.51	0.45	0.533	0.50	0.53	0.472	0.51	0.33	0.638	0.47	0.43	0.005
Oki96158-278	0.795	0.33	0.27	0.611	0.48	0.50	0.432	0.50	0.55	0.355	0.47	0.71	0.475	0.51	0.50	0.115
Oki96222-70	0.607	0.48	0.45	0.667	0.45	0.61	0.512	0.51	0.56	0.408	0.49	0.50	0.817	0.30	0.27	0.088
Oki96376-63	0.557	0.50	0.48	0.389	0.48	0.61	0.844	0.27	0.27	0.750	0.38	0.45	0.659	0.46	0.39	0.120
Oki97660-149	0.227	0.36	0.32	0.419	0.49	0.51	0.044	0.09	0.09	0.145	0.25	0.24	0.061	0.12	0.12	0.152
Oki97954-228	0.500	0.51	0.48	0.395	0.48	0.42	0.600	0.49	0.49	0.615	0.48	0.62	0.500	0.51	0.46	0.018
Oki100771-83	0.023	0.05	0.05	0.000	0.00	0.00	0.136	0.24	0.18	0.041	0.08	0.08	0.000	0.00	0.00	0.098
Oki100884-210	0.000	0.00	0.00	0.000	0.00	0.00	0.000	0.00	0.00	0.000	0.00	0.00	0.024	0.05	0.05	0.013
Oki100974-293	1.000	0.00	0.00	1.000	0.00	0.00	0.933	0.13	0.13	0.872	0.23	0.26	0.854	0.25	0.29	0.052
Oki101119-1006	0.159	0.27	0.32	0.421	0.49	0.53	0.489	0.51	0.36	0.459	0.50	0.54	0.573	0.50	0.51	0.096
Oki101419-103	0.625	0.47	0.48	0.368	0.47	0.47	0.522	0.50	0.56	0.538	0.50	0.46	0.563	0.50	0.38	0.022
Oki101554-359	0.011	0.02	0.02	0.000	0.00	0.00	0.011	0.02	0.02	0.086	0.16	0.17	0.026	0.05	0.05	0.045
Oki101770-525	0.000	0.00	0.00	0.000	0.00	0.00	0.233	0.36	0.33	0.269	0.40	0.44	0.110	0.20	0.17	0.115
Oki102213-604	0.930	0.13	0.14	0.947	0.10	0.11	0.705	0.42	0.41	0.910	0.17	0.18	0.890	0.20	0.22	0.106
Oki102267-166	0.136	0.24	0.18	0.053	0.10	0.05	0.022	0.04	0.04	0.171	0.29	0.29	0.232	0.36	0.37	0.053
Oki102414-499	0.284	0.41	0.39	0.763	0.37	0.37	0.433	0.50	0.56	0.538	0.50	0.62	0.293	0.42	0.44	0.140
Oki102457-67	0.511	0.51	0.57	0.882	0.21	0.24	0.922	0.15	0.16	0.987	0.03	0.03	0.988	0.02	0.02	0.438

Continued on next page

Table 1.4 – continued from previous page

	Nehalem R. (NFH) N=44			**Columbia R. (BH) N=38**			**Quinault R. (QNFH) N=45**			**Green R. (SCH) N=39**			**Fraser R. (CRH) N=41**			
Assay	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	F_{ST}
Oki102801-511	0.045	0.09	0.09	0.026	0.05	0.05	0.167	0.28	0.33	0.042	0.08	0.08	0.049	0.09	0.10	0.064
Oki102867-667	0.000	0.00	0.00	0.000	0.00	0.00	0.000	0.00	0.00	0.000	0.00	0.00	0.000	0.00	0.00	0.000
Oki103271-161	0.580	0.49	0.52	0.750	0.38	0.44	0.822	0.30	0.27	0.833	0.28	0.28	0.963	0.07	0.07	0.120
Oki103577-70	0.442	0.50	0.47	0.324	0.44	0.53	0.305	0.43	0.46	0.410	0.49	0.56	0.188	0.31	0.33	0.035
Oki103713-182	0.136	0.24	0.23	0.434	0.50	0.39	0.211	0.34	0.33	0.218	0.35	0.38	0.293	0.42	0.29	0.054
Oki104515-99	0.750	0.38	0.41	0.357	0.47	0.54	0.656	0.46	0.42	0.500	0.51	0.40	0.232	0.36	0.32	0.185
Oki104519-45	0.636	0.47	0.55	0.500	0.51	0.63	0.750	0.38	0.27	0.842	0.27	0.26	0.805	0.32	0.34	0.078
Oki104569-261	0.000	0.00	0.00	0.000	0.00	0.00	0.000	0.00	0.00	0.000	0.00	0.00	0.061	0.12	0.12	0.050
Oki105105-245	0.034	0.07	0.07	0.066	0.12	0.13	0.378	0.48	0.44	0.382	0.48	0.50	0.256	0.39	0.36	0.137
Oki105115-49	0.170	0.29	0.34	0.216	0.34	0.38	0.278	0.41	0.42	0.244	0.37	0.33	0.256	0.39	0.46	0.000
Oki105132-169	0.547	0.50	0.44	0.581	0.49	0.46	0.364	0.47	0.55	0.471	0.51	0.53	0.338	0.45	0.53	0.036
Oki105235-460	0.263	0.39	0.38	0.279	0.41	0.38	0.730	0.40	0.38	0.712	0.42	0.45	0.818	0.30	0.24	0.271
Oki105385-521	0.352	0.46	0.52	0.153	0.26	0.25	0.144	0.25	0.29	0.068	0.13	0.14	0.073	0.14	0.10	0.112
Oki105407-161	0.591	0.49	0.41	0.500	0.51	0.57	0.651	0.46	0.47	0.700	0.43	0.31	0.598	0.49	0.46	0.007
Oki105897-298	0.705	0.42	0.41	0.878	0.22	0.24	0.739	0.39	0.34	0.857	0.25	0.23	0.863	0.24	0.28	0.030
Oki106172-60	0.261	0.39	0.34	0.371	0.47	0.40	0.278	0.41	0.42	0.372	0.47	0.59	0.415	0.49	0.49	0.009
Oki106313-353	0.511	0.51	0.48	0.447	0.50	0.63	0.727	0.40	0.32	0.462	0.50	0.51	0.439	0.50	0.44	0.051
Oki106419-292	0.307	0.43	0.43	0.289	0.42	0.53	0.267	0.40	0.40	0.167	0.28	0.28	0.317	0.44	0.34	0.005
Oki106479-278	0.989	0.49	0.43	0.500	0.23	0.26	0.500	0.50	0.44	0.500	0.46	0.45	0.500	0.46	0.44	0.104
Oki107336-45	0.702	0.42	0.45	0.568	0.50	0.54	0.411	0.49	0.47	0.378	0.48	0.49	0.313	0.44	0.38	0.091
Oki107607-213	0.068	0.13	0.14	0.284	0.41	0.30	0.182	0.30	0.32	0.092	0.17	0.18	0.134	0.24	0.27	0.044
Oki107974-46	0.420	0.49	0.61	0.270	0.40	0.43	0.386	0.48	0.59	0.368	0.47	0.53	0.375	0.47	0.35	0.001
Oki108505-331	0.884	0.21	0.19	0.882	0.21	0.18	0.933	0.13	0.13	0.934	0.12	0.13	1.000	0.00	0.00	0.016
Oki109243-480	0.057	0.11	0.11	0.145	0.25	0.24	0.100	0.18	0.20	0.064	0.12	0.13	0.122	0.22	0.15	0.002
Oki109525-359	0.013	0.03	0.03	0.233	0.36	0.13*	0.180	0.30	0.12	0.188	0.33	0.13	0.435	0.50	0.09*	0.162

Continued on next page

Table 1.4 – continued from previous page

	Nehalem R. (NFH) N=44			**Columbia R. (BH) N=38**			**Quinault R. (QNFH) N=45**			**Green R. (SCH) N=39**			**Fraser R. (CRH) N=41**			
Assay	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	F_{ST}
Oki109651-152	0.455	0.50	0.68	0.514	0.51	0.36	0.602	0.48	0.48	0.449	0.50	0.59	0.561	0.50	0.49	0.007
Oki109874-122	0.523	0.50	0.50	0.365	0.47	0.46	0.211	0.34	0.29	0.216	0.34	0.27	0.146	0.25	0.20	0.034
Oki109894-418	0.452	0.50	0.52	0.716	0.41	0.51	0.744	0.38	0.33	0.724	0.41	0.39	0.692	0.43	0.56	0.065
Oki110064-418	0.114	0.20	0.23	0.000	0.00	0.00	0.000	0.00	0.00	0.000	0.00	0.00	0.000	0.00	0.00	0.104
Oki110078-191	0.625	0.47	0.52	0.486	0.51	0.59	0.122	0.22	0.24	0.236	0.37	0.31	0.110	0.20	0.22	0.251
Oki110381-77	0.898	0.19	0.20	0.592	0.49	0.55	0.811	0.31	0.20	0.730	0.40	0.54	0.927	0.14	0.15	0.102
Oki110689-43	0.227	0.36	0.27	0.118	0.21	0.24	0.044	0.09	0.09	0.039	0.08	0.08	0.000	0.00	0.00	0.107
Oki111681-407	0.058	0.11	0.12	0.026	0.05	0.05	0.011	0.02	0.02	0.000	0.00	0.00	0.000	0.00	0.00	0.025
Oki113457-324	0.814	0.31	0.28	0.776	0.35	0.29	0.622	0.48	0.49	0.423	0.49	0.44	0.402	0.49	0.46	0.136
Oki113979-170	0.000	0.00	0.00	0.039	0.08	0.08	0.159	0.27	0.27	0.141	0.25	0.23	0.110	0.20	0.22	0.038
Oki114315-360	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	1.000	0.00	0.00	0.000
Oki114448-101	0.977	0.04	0.05	0.730	0.40	0.49	0.651	0.46	0.42	0.654	0.46	0.49	0.707	0.42	0.49	0.082
Oki114587-309	0.398	0.48	0.39	0.316	0.44	0.42	0.389	0.48	0.51	0.316	0.44	0.58	0.402	0.49	0.37	0.000
Oki115987-366	0.000	0.00	0.00	0.013	0.03	0.03	0.011	0.02	0.02	0.013	0.03	0.03	0.000	0.00	0.00	0.000
Oki116362-411	0.205	0.33	0.32	0.595	0.49	0.54	0.378	0.48	0.49	0.397	0.49	0.44	0.671	0.45	0.46	0.131
Oki116865-244	0.523	0.50	0.59	0.671	0.45	0.54	0.826	0.29	0.35	0.611	0.48	0.56	0.488	0.51	0.43	0.072
Oki117043-374	0.170	0.29	0.25	0.068	0.13	0.14	0.200	0.32	0.27	0.197	0.32	0.33	0.195	0.32	0.29	0.005
Oki117144-64	0.761	0.37	0.30	0.500	0.51	0.53	0.768	0.36	0.37	0.824	0.29	0.29	0.622	0.48	0.56	0.063
Oki117286-291	0.000	0.00	0.00	0.000	0.00	0.00	0.000	0.00	0.00	0.013	0.03	0.03	0.000	0.00	0.00	0.001
Oki117742-259	0.372	0.47	0.33	0.186	0.31	0.26	0.122	0.22	0.20	0.145	0.25	0.29	0.110	0.20	0.17	0.083
Oki117815-369	0.080	0.15	0.16	0.095	0.17	0.19	0.044	0.09	0.09	0.206	0.33	0.35	0.303	0.43	0.45	0.091
Oki118152-314	0.378	0.48	0.46	0.514	0.51	0.53	0.478	0.50	0.64	0.635	0.47	0.46	0.795	0.33	0.36	0.093
Oki118175-264	0.455	0.50	0.45	0.516	0.51	0.58	0.686	0.44	0.53	0.526	0.51	0.59	0.263	0.39	0.33	0.090
Oki118654-330	0.477	0.50	0.59	0.284	0.41	0.41	0.511	0.51	0.52	0.487	0.51	0.34	0.650	0.46	0.45	0.050
Oki119108-320	0.068	0.13	0.14	0.079	0.15	0.16	0.011	0.02	0.02	0.000	0.00	0.00	0.000	0.00	0.00	0.029

Continued on next page

Table 1.4 – continued from previous page

	Nehalem R. (NFH) N=44			**Columbia R. (BH) N=38**			**Quinault R. (QNFH) N=45**			**Green R. (SCH) N=39**			**Fraser R. (CRH) N=41**			
Assay	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	AF	H_E	H_O	F_{ST}
Oki120024-226	0.419	0.49	0.42	0.724	0.41	0.34	0.667	0.45	0.49	0.603	0.49	0.54	0.646	0.46	0.46	0.049
Oki121006-412	0.163	0.28	0.33	0.176	0.29	0.35	0.136	0.24	0.27	0.103	0.19	0.21	0.038	0.07	0.08	0.014
Oki122593-430	0.767	0.36	0.33	0.855	0.25	0.24	0.800	0.32	0.36	0.818	0.30	0.30	0.976	0.05	0.05	0.033
Oki123044-68	0.432	0.50	0.59	0.333	0.45	0.36	0.238	0.37	0.33	0.410	0.49	0.51	0.317	0.44	0.44	0.017
Oki123205-88	0.128	0.23	0.26	0.026	0.05	0.05	0.167	0.28	0.29	0.346	0.46	0.54	0.138	0.24	0.28	0.090
Oki123470-92	1.000	0.00	0.00	0.895	0.19	0.11	0.989	0.02	0.02	0.943	0.11	0.06	1.000	0.00	0.00	0.050
Oki123921-90	0.093	0.17	0.14	0.132	0.23	0.26	0.033	0.07	0.07	0.026	0.05	0.05	0.037	0.07	0.07	0.024
Oki124162-62	0.595	0.49	0.27	0.829	0.29	0.29	0.578	0.49	0.53	0.551	0.50	0.44	0.592	0.49	0.45	0.037
Oki125998-340	0.186	0.31	0.37	0.149	0.26	0.30	0.189	0.31	0.33	0.026	0.05	0.05	0.085	0.16	0.17	0.029
Oki126160-142	0.375	0.47	0.52	0.355	0.46	0.50	0.333	0.45	0.36	0.141	0.25	0.23	0.098	0.18	0.20	0.069
Oki127236-383	0.071	0.13	0.14	0.039	0.08	0.08	0.056	0.11	0.11	0.359	0.47	0.46	0.122	0.22	0.20	0.180
Oki127760-301	0.091	0.17	0.18	0.145	0.25	0.18	0.011	0.02	0.02	0.000	0.00	0.00	0.000	0.00	0.00	0.068
Oki128302-547	0.047	0.09	0.09	0.105	0.19	0.16	0.114	0.20	0.14	0.066	0.12	0.13	0.122	0.22	0.10	0.000
Oki128693-70	0.352	0.46	0.61	0.095	0.17	0.19	0.400	0.49	0.53	0.171	0.29	0.29	0.100	0.18	0.20	0.108
Oki128757-232	0.000	0.00	0.00	0.013	0.03	0.03	0.000	0.00	0.00	0.000	0.00	0.00	0.037	0.07	0.07	0.021
Oki128851-185	0.511	0.51	0.43	0.434	0.50	0.55	0.333	0.45	0.40	0.321	0.44	0.49	0.372	0.47	0.33	0.015
Oki129870-552	0.300	0.43	0.15*	0.132	0.23	0.26	0.133	0.23	0.22	0.365	0.47	0.46	0.225	0.35	0.35	0.044
Oki130295-48	0.648	0.46	0.39	0.868	0.23	0.21	0.978	0.04	0.04	0.974	0.05	0.05	0.951	0.09	0.10	0.239
Oki130524-184	0.512	0.51	0.45	0.347	0.46	0.64	0.611	0.48	0.47	0.403	0.49	0.58	0.207	0.33	0.41	0.096
Oki131147-353	0.875	0.22	0.16	0.643	0.47	0.43	0.489	0.51	0.44	0.270	0.40	0.38	0.700	0.43	0.50	0.206
Oki131460-243	0.261	0.39	0.39	0.487	0.51	0.61	0.467	0.50	0.49	0.500	0.51	0.35	0.500	0.51	0.46	0.033
Oki131802-368	0.011	0.02	0.02	0.054	0.10	0.05	0.193	0.32	0.34	0.167	0.28	0.23	0.366	0.47	0.49	0.138
Mean	0.391	0.30	0.30	0.385	0.30	0.31	0.398	0.31	0.30	0.386	0.31	0.31	0.388	0.28	0.27	0.073
Polymorphic Loci (%)		87.0			89.8			93.5			91.7			87.4		

Preliminary BLAST (Basic Local Alignment Search Tool, NCBI) results and annotation of the target SNP can be found in Table 1.5. Annotation has been included for the loci described in this chapter (References 1), as well as an additional 17 loci (References 2 and 3) that make up the genotyping panel in Chapter 2. Out of the 109 SNPs, 76 were found in exons, while 30 were found in introns based on alignment of the consensus with the EST sequences. A total of 74 consensus sequences matched a known gene from GeneBank with E-values ranging from 9E-146 to zero (smaller numbers indicating higher similarity). Eighteen SNPs were identified in either the 5' or 3' untranslated region (UTR), while for five loci, annotated translation (n.t.) was unavailable. The introns for 15 loci were found within the coding sequences (CDS) of a gene, while others were found upstream or downstream of the annotated fragment. Of the variation found in CDS exons, three were synonymous substitutions while five were non-synonymous substitutions.

Table 1.5: Preliminary BLAST results (BLAST hit and e-value) and annotation of the target SNP for the loci described here (Reference 1) and for an additional 17 loci (References 2 and 3) that are part of the final genotyping panel described in Chapter 2. Also included is whether the variation is present in an intron or exon and its location with respect to the described gene, either in coding sequence (CDS) or untranslated regions (UTR). No translation (n.t.) was available for 5 loci. For CDS exons, a single amino acid is indicated for synonymous substitutions, while both amino acids are included for non-synonymous substitutions. Reference codes are as follows: 1. Starks Chapter 1; 2. Smith *et al.* 2006; 3. Campbell and Narum 2011.

Assay name	**Ref**	**SNP**	**Genic location**	**BLAST # [E-value]**	**Description**
Oki94903-192	1	exon			
Oki95318-100	1	exon			
Oki96127-66	1	exon	3' UTR	BT047774 [0]	*Salmo salar* Thioredoxin domain-containing protein
Oki96158-278	1	exon	CDS 1 [Pro]	AY190723 [1.40E-29]	*Pagrus major* RAP2B-like protein
Oki96222-70	1	intron	CDS	AB326306 [2.02E-45]	*Solea senegalensis* elongation factor 1 alpha isoform
Oki96376-63	1	intron			
Oki97660-149	1	exon	5' of gene	EU025717 [7.07E-29]	*Salmo salar* single-strand selective monofunctional uracil
Oki97954-228	1	intron	CDS	EU325855 [2.10E-96]	*O. masou* formosanus retinol-binding protein 2
Oki100771-83	1	exon			
Oki100884-210	1	exon	3' UTR	NM_001140998 [7.34E-173]	*Salmo salar* centrosomal protein 97
Oki100974-293	1	exon			
Oki101119-1006	1	intron			
Oki101419-103	1	intron	CDS	BT072612 [0]	*Salmo salar* Alpha-N-acetylgalactosaminidase
Oki101554-359	1	exon	3' UTR	BT058832 [8.48E-177]	*S. salar* NMDA receptor-regulated protein 1
Oki101770-525	1	intron	CDS	BT078766 [6.96E-86]	*Esox lucius* Cartilage-associated protein precursor
Oki102213-604	1	intron	CDS	BT073541 [5.62E-76]	*O. mykiss* Leukocyte cell-derived chemotaxin 2 precursor
Oki102267-166	1	exon	n.t.	EU621898 [8.77E-28]	*Salmo salar* interferon alpha 1-like gene
Oki102414-499	1	exon			
Oki102457-67	1	intron		NM_001124228 [2.41E-64]	*O. mykiss* heat shock protein 70a
Oki102801-511	1	exon			
Oki102867-667	1	intron			
Oki103271-161	1	exon		EU621901 [1.82E-31]	*O. tshawytscha* growth hormone 2 gene
Oki103577-70	1	intron	CDS	BT058741 [2.52E-45]	*S. salar* NADPH–cytochrome P450 reductase
Oki103713-182	1	exon			
Oki104515-99	1	intron			
Oki104519-45	1	exon	3' UTR	NM_001173633 [1.30E-34]	*S. salar* Nucleolar transcription factor 1
Oki104569-261	1	exon			
Oki105105-245	1	exon	CDS [Arg>Ser]	AF483538 [9.12E-90]	*O. mykiss* VHSV-induced protein
Oki105115-49	1	exon			
Oki105132-169	1	exon	3' of gene	NM_001173959 [1.33E-139]	*S. salar* RAP2C, member of RAS oncogene family

Continued on next page

Table 1.5 – continued from previous page

Assay name	Ref	SNP	Genic location	BLAST # [E-value]	Description
Oki105235-460	1	exon	CDS (poor match)	BT057774 [8.67E-105]	*S. salar* Growth arrest and DNA-damage-inducible protein
Oki105385-521	1	exon			
Oki105407-161	1	exon	5' of gene	BT047755 [8.13E-133]	*S. salar* 60S ribosomal protein L36a
Oki105897-298	1	exon	3' UTR	NM_001173890 [2.35E-127]	*S. salar* Uridine 5-monophosphate synthase
Oki106172-60	1	exon			
Oki106313-353	1	intron	CDS	NM_001141707 [9.98E-129]	*S. salar* Wilms tumor 1 associated protein
Oki106419-292	1	intron	3' of gene	AF281332 [5.35E-77]	*O. mykiss* biotinidase fragment 1
Oki106479-278	1	exon	n.t. (pseudogene)	BT072720 [0]	*S. salar* Transformer-2 protein homolog
Oki107336-45	1	exon	CDS [Gly>Val]	EU816603 [0]	*S. salar* (NUDT6)
Oki107607-213	1	exon			
Oki107974-46	1	exon	3' UTR	EU025714 [0]	*S. salar* retinoic acid receptor gamma a
Oki108505-331	1	intron			
Oki109243-480	1	exon	3' UTR	NM_001140063 [2.15E-59]	*S. salar* Kunitz-type protease inhibitor 2
Oki109525-359	1	intron	CDS	NM_001124667 [1.46E-128]	*O. mykiss* prostaglandin-endoperoxide synthase 2b
Oki109651-152	1	intron	CDS (poor match)	NM_001124398 [5.65E-46]	*O. mykiss* secreted phosphoprotein 24
Oki109874-122	1	exon	n.t.	ET358838 [4.43E-27]	*S. salar* CHORI-214
Oki109894-418	1	intron			
Oki110064-418	1	exon	CDS [Leu>Pro]	NM_001165121 [4.29E-73]	*O. mykiss* lipopolysaccharide-induced TNF factor
Oki110078-191	1	exon			
Oki110381-77	1	exon			
Oki110689-43	1	exon			
Oki111681-407	1	exon	CDS [Tyr>Phe]	NM_001124224 [0]	*O. mykiss* NK2 homeobox 1b
Oki113457-324	1	exon	n.t. (pseudogene)	BT071894 [3.63E-85]	*S. salar* C-ets-2
Oki113979-170	1	exon	3' UTR	BT073942 [5.66E-92]	*O. mykiss* Leptin receptor overlapping transcript-like 1
Oki114315-360	1	exon	3' UTR	BT048189 [0]	*S. salar* Tumor protein D54
Oki114448-101	1	exon			
Oki114587-309	1	intron	3' of gene (poor match)	NM_001165381 [9.40E-42]	*S. salar* v-crk sarcoma virus CT10 oncogene homolog
Oki115987-366	1	intron	3' of gene (poor match)	BT047671 [2.67E-174]	*S. salar* canopy homolog 2 precursor
Oki116362-411	1	intron	CDS (poor match)	BT058381 [4.47E-47]	*S. salar* adhesion regulating molecule 1
Oki116865-244	1	exon	3' of gene	NM_001140533 [4.13E-20]	*S. salar* Guanylate kinase
Oki117043-374	1	exon			
Oki117144-64	1	exon	n.t.	BT072411 [0]	*S. salar* Transmembrane and ubiquitin-like protein 2
Oki117286-291	1	exon			
Oki117742-259	1	exon	CDS [His]	BT045072 [0]	*S. salar* CREB/ATF bZIP transcription factor
Oki117815-369	1	exon	3' UTR	NM_001140301 [8.56E-167]	*S. salar* glutaredoxin (thioltransferase)
Oki118152-314	1	intron			
Oki118175-264	1	exon			
Oki118654-330	1	exon			

Continued on next page

Table 1.5 – continued from previous page

Assay name	Ref	SNP	Genic location	BLAST # [E-value]	Description
Oki119108-320	1	exon	CDS [Ala]	NM_001146421 [1.08E-99]	*S. salar* CK046 protein
Oki120024-226	1	intron	CDS	BT072200 [0]	*S. salar* Solute carrier family 28 member 3
Oki120255-113-sppID	1	intron	CDS (poor match)	BT058586 [1.82E-115]	*S. salar* C6orf64
Oki121006-412	1	exon	CDS [Leu>Ile]	BT049490 [0]	*S. salar* Immediate early response gene 5 protein
Oki122593-430	1	intron	CDS	BT046018 [0]	*S. salar* Voltage-dep. anion-selective channel protein 2
Oki123044-68	1	intron	CDS	NM_001173712 [2.12E-26]	*S. salar* Phosphatidylinositol-binding clathrin assem. protn.
Oki123205-88	1	exon	3' UTR	NM_001139756 [6.84E-38]	*S. salar* transposase-like
Oki123470-92	1	exon	3' UTR	BT059490 [0]	*S. salar* Kunitz-type protease inhibitor 1 precursor
Oki123921-90	1	intron			
Oki124162-62	1	intron			
Oki125998-340	1	exon			
Oki126160-142	1	exon			
Oki127236-383	1	exon			
Oki127760-301	1	exon	CDS (poor match)	NM_001140111 [6.59E-101]	*S. salar* Zinc finger protein 503
Oki128302-547	1	intron	CDS (poor match)	BT073572 [5.14E-138]	*O. mykiss* 39S ribosomal protein L20
Oki128693-70	1	exon	CDS (poor match)	NM_001141075 [3.57E-94]	*S. salar* stathmin-like 4 (stmn4)
Oki128757-232	1	exon	3' UTR	BT057575 [3.08E-176]	*S. salar* Thymosin beta-11
Oki128851-185	1	exon	5' UTR	NM_001124601 [1.86E-48]	*O. mykiss* differentially regulated trout protein 1
Oki129870-552	1	intron			
Oki130295-48	1	exon	3' UTR	NM_001124734 [0]	*O. mykiss* simple type II keratin K8b (S2)
Oki130524-184	1	exon			
Oki131147-353	1	exon	3' UTR	NM_001139629 [0]	*S. salar* high-mobility group box 1
Oki131460-243	1	intron	CDS (poor match)	BT072067 [5.46E-74]	*S. salar* Neural cell adhesion mol. L1-like protn. prec.
Oki131802-368	1	exon			
Oki_afp4-10	3	exon		FN396363 [1E-129]	*O. mykiss* Type-4 ice-structuring protein precursor
Oki_arp-105	3	unk.	CDS (partial)	EU682504 [0]	*O. mykiss* acidic ribosomal phosphoprotein gene
Oki_carban-140	3	exon		NM_001124220 [0]	*O. mykiss* carbonic anhydrase 1
Oki_gdh-189	3	exon	3' UTR	EF042600 [0]	*O. tshawytscha* glutamate dehydrogenase gene
Oki_gh-183	3	unk.	CDS	U14551 [1E-160]	*O. nerka* type-1 growth hormone gene
Oki_gshpx-152	3	exon	CDS (partial)	AF281338 [7E-150]	*O. mykiss* glutathione peroxidase
Oki_HGFA-311	2	exon	CDS (partial)	AF281356 [4E-121]	*O. mykiss* putative hepatocyte growth factor activator
Oki_hsc713-56	3	exon	3' UTR	FJ772769 [0]	*O. mykiss* isolate Omy_29 heat shock cognate 71 gene
Oki_ins-167	2	unk.		X13559 [0]	*Oncorhynchus* insulin gene for preproinsulin
Oki_itpa-85	3	exon	CDS (pseudogene)	BT071971 [9E-146]	*Salmo salar* clone ssal-rgf-505-256
Oki_LWSop-554	2	exon	CDS	AY214145 [2E-151]	*O. kisutch* LWS opsin mRNA
Oki_nips-159	3	exon		NM_001160597 [3E-136]	*O. mykiss* NipSnap2 (nips2)
Oki_p53-20	3	exon		NM_001124692 [3E-121]	*O. mykiss* tumor protein p53 (tp53)

Continued on next page

Table 1.5 – continued from previous page

Assay name	**Ref**	**SNP**	**Genic location**	**BLAST # [E-value]**	**Description**
Oki_pigh-33	3	exon		NM_001160586 [3E-161]	*O. mykiss* phosphatidylinositol glycan anchor biosynthesis
Oki_rpo2j-235	3	exon		NM_001160575 [8E-157]	*O. mykiss* DNA-directed RNA polymerase II subunit J
Oki_SClkF2R2-120	2	exon	CDS (partial)	DQ780892 [9E-105]	*O. tshawytscha* CLOCK1a gene
Oki_txnip-35	3	exon	CDS	BT058954 [8E-80]	*S. salar* Thioredoxin-interacting protein, class H

1.4 Discussion

Presented is a collection of original markers designed for 91 single nucleotide polymorphisms in coho salmon, an economically valuable Pacific salmonid species facing extinction in much of its native range, as well as one diagnostic assay that distinguishes coho and Chinook salmon. I show that by utilizing a public database of existing *O. mykiss* ESTs, high volume SNP discovery in a related non-model organism is successful. I performed an extensive sequencing effort that assessed 234 gene fragments with a total consensus length of 125 kb in an average of 17 individual salmon. This effort has also significantly increased the amount of SNP loci currently available for coho salmon. The 91 SNP markers reported here are broadly polymorphic and should provide a valuable genetic resource for purposes such as genetic stock identification (GSI), individual identification, pedigree reconstruction, phylogeography, and ecological applications. The species diagnostic assay will also be useful in identifying coho salmon accidentally and/or illegally caught in the Chinook fishery, as well as putative Chinook salmon samples collected in coho salmon sampling efforts.

The development and use of molecular markers for the purpose of estimating stock origin of fish in salmon fisheries is becoming a more widespread option for fishery management. For GSI, establishment of a reference baseline containing markers with appropriate allelic frequency differences between populations is required in order to estimate stock proportions from a mixed fishery sample (Garza and Anderson 2007, Seeb *et al.* 2007). Because of the bi-allelic properties of SNPs, large numbers of informative

SNP loci are needed for accurate assignment. It has also been suggested that the inclusion of several SNP loci that have high F_{ST} values (under divergent selection), may also increase the accuracy of GSI assignments (Hess *et al.* 2011). The mean F_{ST} estimated for the 91 SNP loci presented here was 0.131 over the ten populations, and individual pairwise F_{ST} estimates ranged from 0.006-0.445, suggesting that some loci are under greater diversifying selective pressures and might be more informative than others for GSI purposes (Nielsen 2005). Progress towards the establishment of an informative baseline of SNP population markers has begun in recent years (Seeb and Seeb 2006) and the 91 SNP assays described here should add significantly to available genetic resources for coho salmon. Likewise, preliminary analyses suggest that assignment accuracy for GSI in the validated populations is similar to that retrieved from the 18 microsatellite loci currently in use for coho salmon at the NOAA Santa Cruz laboratory (Bjorkstedt *et al.* 2005) when previously designed markers (Smith *et al.* 2006, Campbell and Narum 2011) were combined with the most informative assays from the present study (data not shown).

Additionally, the SNP loci presented here will make intergenerational genetic tagging possible through large-scale parentage inference because 80-100 SNP loci are needed in order to minimize tag recovery error rates (Anderson and Garza 2006). Parentage-based tagging (PBT) is largely centered on the idea that by genotyping parental fish, either hatchery broodstock or wild spawners, genetic tags from their offspring can be recovered and used to provide information on their stock and cohort of origin (age) of subsequent fish (Anderson and Garza 2006, Garza and Anderson 2007).

The large number of polymorphic SNP loci specifically designed for coho salmon in this study, as well as the few others currently available (Smith *et al.* 2005, Smith *et al.* 2006, Campbell and Narum 2011) will help to provide a better understanding of coho life history traits such as age at maturation and effective population size, reproductive success, mortality rates, and migration rates, both in natural populations and in hatchery settings.

Ultimately, SNP markers will help to inform decisions in species conservation and hatchery management. Unlike microsatellites which are based on fragment size, SNP markers are directly based on DNA sequences. This allows for resulting SNP data from different labs around the world to be highly portable and easy to combine. Such characteristics reduce the time and cost that is incurred in the generation and standardization of microsatellite data. SNP loci also have the potential to provide more reliable data due to a low rate of mutation, an occurrence that microsatellite loci are frequently subjected to at high rates (Picoult-Newberg *et al.* 1999). As such, the growing number of informative SNP loci becoming available to researchers will be extremely valuable for future uses.

This SNP discovery effort used traditional Sanger sequencing methods to sequence PCR products from 275 ESTs from an *O. mykiss* database, and the results of this effort were quite successful. The ascertainment panel included samples from several populations of coho salmon representing a large portion of the species' natural range, and specifically included representatives of California populations that are currently being studied in this lab. The resulting assays produced in this research effort will be

employed in future studies and monitoring programs involving these populations. Only assays for which all three genotypes were observed were designed, regardless of which individuals or population representatives harbored variation. In doing so, markers were chosen that had higher mean minor allele frequencies (0.266-0.398) as well as an increased likelihood of being more generally useful in the species. This means however that the minor allele frequency results are upwardly biased due to the under representation of rare alleles. The ascertainment and validation panel also included fish with generally small effective population sizes in the Russian River, largely due to recent and severe reductions in population size. The inclusion of these fish is presumably what caused the high instance of apparent linkage disequilibrium, as the analysis showed that almost all significant linkage disequilibrium occurred in locus pairs within the Russian River (data not shown) even though all loci were designed from separate genes. The under-representation of rare alleles may be even more pronounced in this population due to the potential loss of genetic variation that generally occurs when populations have been greatly reduced (Garza and Williamson 2001).

The ascertainment panel was designed to locate variation in a diverse group of coho salmon in order to reduce ascertainment bias. However it was not possible to completely eliminate ascertainment bias due to the extent of phylogeographic diversity within the species and the large amount of additional sequencing effort that this would entail. The need for more informative SNP assays for all populations of coho salmon is ever increasing. As a result, more recent efforts are turning to next generation sequencing (NGS) methods such as pyrosequencing for SNP discovery in salmonids.

These methods have the potential to produce a large amount of genetic sequence data that harbor a large number of polymorphic sites. Sánchez *et al.* (2009) produced close to 130,000 contigs in *O. mykiss* that contained 20,000 putative SNPs. However their resulting conversion rate upon validation was very low (48%) when compared to our conversion rate (80%) because NGS is apparently unable to differentiate between true SNP sites and sequencing artifacts. Certain protocols are available, however, that might improve NGS results (Campbell and Narum 2011). Traditional Sanger sequencing of targeted EST loci has its advantages, in that the functional roles for certain genes are often already known, and it is easier to sequence the same genomic regions across a diverse panel of individuals. In this study, only a small portion (275) of the ESTs from the *O. mykiss* Gene Index, in which there are nearly 100 kb of EST sequence available for exploration, were examined. Either through another targeted EST sequencing effort or through the use of NGS technology, an additional sequencing and SNP discovery effort will obviously need to take place in order to further reduce ascertainment bias and discover more coho salmon variation.

Chapter 2

The use of SNPs for intergenerational genetic tagging allows pedigree-based inference in a hatchery population of coho salmon (*Oncorhynchus kisutch*)

2.1 Introduction

Mark and recapture methods have proven to be very useful for studying the ecology, population biology and evolutionary history of organisms over time. Mark-recapture, or physical tagging, experiments have been performed in many animal groups including mammals (Herrera 1992; Rogers *et al.* 1996), birds (Morton 1992; Johnston *et al.* 1997), amphibians (Berven and Grudzien 1990), insects (Weslien and Lindelow 1990; Sutcliffe *et al.* 1997) and fish (Miller and Simenstad 1997; Trajano 1997), and they

have provided a powerful way by which to estimate population size and distributions, migration patterns, responses to varying environmental parameters and effects of natural selection. However, this method is not always a viable option for certain organisms because tag attachment and recovery can be too invasive, tags are often expensive and/or prone to loss, the animals themselves are too phenotypically similar to be easily distinguished in the wild, and actual tag recovery rates can be low in organisms that experience high mortality (Palsbøll *et al.* 1997; Woods *et al.* 1999; Hankin *et al.* 2005).

For these reasons, many researchers have turned to genetic tagging as an alternative to the traditional mark-recapture methods. Several studies have demonstrated the utility of using genetics as a means of tagging whales (Palsbøll *et al.* 1997) and bears (Woods *et al.* 1999), as well as bush rats (Peakall *et al.* 2006), fur seals (Hoffman *et al.* 2006), and salamanders (Unger *et al.* 2012). In these cases, an individual's DNA or genotype serves as the tag. The same individual is then 'recaptured' when its matching genotype has been identified in a subsequent sampling effort. In addition to mark-recapture methods, genetic tagging is also employed in order to infer familial relationships, like parent-offspring pairs and siblings. Through these methods, it is possible to estimate certain life history traits and population parameters like effective population size, reproductive success and reproductive strategy, dispersal, fitness and hybridization in both wild and captive groups of individuals (Fowler *et al.* 1998; Blouin 2003; DeWoody 2005; Taylor *et al.* 2007; Wang *et al.* 2008). In one such study, parentage analyses were used to confirm that wildlife crossings prevented population isolation in both grizzly bears and black bears by allowing for continued gene flow when said

populations suffered from habitat fragmentation (Sawaya *et al.* 2014).

Of late, single nucleotide polymorphisms (SNPs) have become an increasingly popular molecular tool for use in population biology. Anderson and Garza (2006) demonstrated that a relatively small number of SNPs (< 100) are necessary in order to provide powerful and accurate parentage inference in large data sets involving highly fecund organisms. Additionally, other benefits associated with SNPs such as their ease of portability, low error rates and suitability for high-throughput genotyping, has led to the development of many new SNP markers (Abadía-Cardoso *et al.* 2011; Clemento *et al.* 2011; Starks Chapter 1). This, in combination with new parentage algorithms/software (Anderson and Garza, 2006; Anderson 2012) designed to quickly and efficiently handle large amounts of SNP data, has made large-scale parentage analysis and pedigree reconstruction in salmonids a reality (Abadía-Cardoso *et al.* 2013; Steele *et al.* 2013).

Coho salmon (*Oncorhynchus kisutch*) is a species of Pacific salmonid, highly prized in both the sport and commercial fishing industries. However, like other salmonids, coho populations residing in the more southern reaches of the species' native range (i.e. California, Oregon and Washington) are rapidly declining or have already been driven to extinction (Frissel 1993; Brown *et al.* 1994; Weitkamp *et al.* 1995). As a result, most coho populations in the contiguous United States are now protected under the Endangered Species Act (ESA). From California to the Pacific Northwest, coho have been divided into seven Evolutionarily Significant Units (ESUs) based on population structure and, of these seven ESUs, five are listed as either endangered, threatened, or a species of

concern under the ESA (FedReg 1995, 1997, 2005). This rapid, dramatic decline has primarily been attributed to freshwater habitat loss resulting from the construction of dams, logging and agriculture, combined with over-fishing and fluctuating ocean conditions (Hare *et al.* 1999; Ruckelshaus *et al.* 2002). As a result, many hatcheries have been established in order to mitigate these losses through supplementation by artificial propagation.

Although salmon hatcheries along the Pacific West Coast of the US produce millions of coho annually, these fish, like other hatchery fish, generally carry the stigma of having a potentially harmful influence on their wild, naturally spawning counterparts. Reduced fitness owing to relaxed selective pressures associated with the hatchery environment, combined with elevated levels of inbreeding, have been implicated as sources for these fitness reductions (Christie *et al.* 2012; Christie *et al.* 2013). As these hatchery fish are released into the wild as juveniles, they then carry the risk of imparting, through introgression, their hatchery-selected genes into wild stocks, resulting in reduced fitness and maladaptation of wild populations (Araki *et al.* 2007; Buhle *et al.* 2009; Quiñones *et al.* 2014).

Ultimately, improvements in hatchery practices are needed in order to increase the genetic variability and overall fitness of these populations, so that the process of conservation and in some cases rehabilitation of wild stocks can begin. In order do this in an effective and beneficial manner, it is first necessary to gain a full understanding of the life history variation and reproductive strategies specific to the populations in question. Although the basic lifecycle of coho salmon is relatively invariable in comparison to

other salmonids like steelhead (*O. mykiss*), the species still displays minor variability in regard to adult sexual maturity and run timing. As coho salmon are anadromous, they can generally spend 1-2 years in fresh water after hatching, and another 1-2 years in the ocean foraging before they migrate back to their natal streams to spawn. Coho salmon are semelparous, meaning that they die after a single spawning event. Most coho salmon along the Pacific west coast reenter freshwater in October when fall rains increase river flow, spawning from November to December and sometimes into January (Weitkamp *et al.* 1995). However, the mouths of many small coastal California streams are obstructed by sandbars for most of the year until sufficient winter rains allow them to be breached, so river reentry and actual spawn timing can be delayed upwards of three months due to delayed rain fall and/or drought (Weitkamp *et al.* 1995; Lestelle 2007).

Fully grasping the combined biological/heritable and environmental mechanisms responsible for shaping these different life history strategies is difficult to do in a highly fecund species that undergoes extensive migrations. In order to attempt to do so, the ability to accurately identify and track individuals, and ultimately families, within and among populations and over generations is both desirable and necessary for meaningful biological inference. Currently, coho salmon are identified as being of hatchery origin by way of physical markings (adipose and maxillary fin clips) and/or by the employment of coded wire tags (CWTs; Hankin *et al.* 2005). CWTs enable hatchery managers to identify ocean caught fish down to the cohort and hatchery stock origin levels, and they provide inference on ocean stock distributions and age specific

mortality. However, only a fraction of individuals per cohort receive a CWT, identification to family level is not possible, and tag retrieval is lethal. In a species that is ESA-listed (even hatchery origin fish) over much of its native range within the contiguous United States, these tags have minimal utility owing in part to the fact that these fish cannot be legally harvested in the ocean, and because injuries resulting from the tagging/marking procedure in itself may lead to increased straying, disease prevalence, and overall mortality in later life stages (Morrison and Zajac 1987; Habicht *et al.* 1998; Elliott and Pascho 2001; Crozier and Kennedy 2002).

The limited tagging capabilities described above leave much to be desired in terms of providing the information needed to develop sound conservation and management guidelines specific to these dwindling populations of coho salmon. As a result, the stage has been set to test and validate the accuracy and effectiveness of large-scale SNP-based parentage inference, or pedigree-based tagging (Anderson and Garza 2006). The rationale behind this tagging method is that a hatchery cohort is tagged by collecting the parental genotypes of the broodstock. Tag retrieval is then performed when the corresponding adult offspring return to the hatchery to spawn; through non-lethal genetic sampling, followed by high confidence pedigree inference (Anderson 2012), individual offspring are then assigned back to their true parent-pairs. Because each parent-pair is likely to produce numerous juvenile offspring that will experience high mortality (i.e. lost tags), being able to assign tags based on parental genotypes rather than physical tag recovery is an efficient method. Additionally, stock origin and cohort specific information, as well as date of birth and reproductive success of certain individuals can

also be determined. As such, this method harbors wide-ranging potential for the evaluation and continued monitoring of salmonid populations, as well as other high fecundity species with complex lifecycles.

This study describes a large-scale, intergenerational genetic tagging experiment in a medium sized hatchery population of coho salmon from the Klamath River, CA, USA. In this experiment, I use a panel of SNP markers developed specifically for coho salmon (Smith *et al.* 2006; Campbell and Narum 2011; Starks Chapter 1) in order to reconstruct with high confidence, individual pedigrees of adults having undergone ocean migrations. Assignment accuracy was confirmed by comparing assignments to hatchery recorded spawn pair records as well as to records of physical hatchery marks assigned to individual fish. The reconstructed parent-offspring trios were then used to evaluate variability in the age structure of offspring cohorts as well as of parental return years, and the relative reproductive success of spawning broodstock. Phenotypic data collected for most parents and their offspring allowed for heritability estimates of length-at-maturity, as well as correlations between female length and the number of their offspring returning to spawn.

Additionally, I performed an in depth comparison of the overall performance of both SNP and microsatellite markers at calculating relatedness. Because prior genetic analyses employing microsatellites and the relatedness statistic, Rxy (Queller and Goodnight 1989), determined that inbreeding was a significant problem in this population (HSRG 2012 per C. Garza draft analysis 2010), genetic spawning matrices are now employed in order to guide mating protocols at the hatchery. These matrices determine

preferred mates based on the degree of relatedness inferred between potential spawners in an effort to reduce inbreeding. As genetically based spawn matrices are to be implemented in all future brood years, and because the SNP panel used in this study was developed with the intention of replacing the microsatellite panel currently in use at the hatchery, it is important to determine whether SNP-generated relatedness values are comparable to those inferred from microsatellites. As such, estimates of relatedness in the spawning populations were compared between the two marker types, and estimates of relatedness in pedigree-confirmed siblings were compared. The effect of parental relatedness on their reproductive success was also assessed using both marker types for comparison. I demonstrate that pedigree-based tagging provides a powerful means for identifying and understanding many biological and population level aspects of an ESA listed species with significant conservation concerns.

2.2 Methods

2.2.1 Study Site

The headwaters of the Klamath River begin in Southern Oregon and flow through Northern California to empty into the Pacific Ocean approximately 32 km south of Crescent City, CA (Figure 2.1). The Klamath is the second largest river in California, with an average discharge of 481 m^3/s and it supports populations of coho salmon, Chinook salmon and steelhead trout. The construction of Copco Dam 1 in 1918 and Iron Gate Dam in 1962 completely obstructed salmonid migration into the upper

reaches of the Klamath Basin. In 1963, Iron Gate Hatchery (IGH) was built at the base of Iron Gate Dam, and in 1966 artificial propagation of coho salmon began in order to mitigate for the loss of spawning and rearing habitat imposed by these barriers.

Coho salmon run time in the Klamath is relatively short, with adults returning to the river from late October to early January. Historically, hatchery staff randomly selected and spawned adult fish returning to Iron Gate at a male to female ratio of 1:1 (generally) until desired egg production was achieved (IGH reports 2004-2010: Hampton 2005; Richey 2006; Chesney 2007; Chesney 2009; Chesney and Knechtle 2011). A size cutoff has traditionally been employed in order to exclude precocious age-two males or "jacks" from broodstock, with the cutoff ranging from 510-580 mm [males 510 mm considered precocious in 2004 (Hampton 2005) and males 580 mm considered precocious in 2008 and 2010 (Chesney 2009; Chesney and Knechtle 2011)]. Prior to 2010, all hatchery marked fish that returned to the hatchery were sacrificed, however in 2010 only spawned individuals were sacrificed while those that were not spawned in the hatchery were released back into the river. Eggs were incubated at the hatchery and juveniles received a left maxillary (LM) bone clip for visual determination of IG hatchery origin prior to release at age one. There are two other coho hatchery programs near the Klamath River: Trinity River Fish Hatchery (TRH) at the base of Lewiston Dam on the Trinity River, a southern tributary of the Klamath, and Cole M. Rivers Hatchery (CRH) which is situated at the base of Lost Creek Dam on the Rogue River. Juvenile coho released from TRH receive a right maxillary (RM) bone clip while those released from CRH are marked by the removal of their adipose (AD) fin. Individuals from both

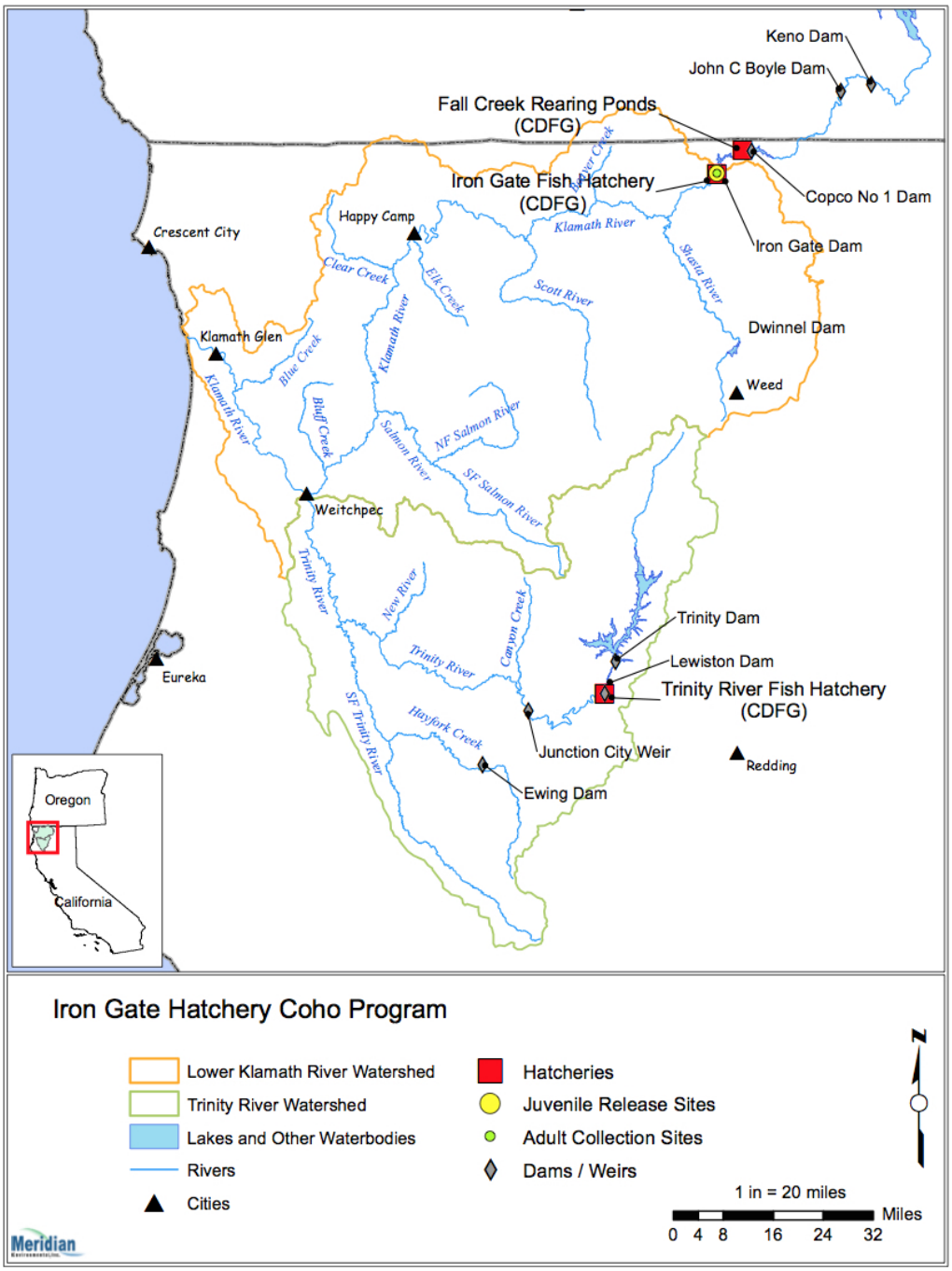

Figure 2.1: Location of Iron Gate Hatchery (IGH) at the base of Iron Gate Dam, on the Klamath River, as well as Trinity River Fish Hatchery (TRH) at the base of the Lewiston Dam. Map originally published in HSRG 2012.

of these two hatcheries frequently return to Iron Gate and prior to 2010 have been incorporated into broodstock (HSRG 2012; Table 2.1).

Table 2.1: Summary of sampling and genotyping effort at Iron Gate Hatchery, Hornbrook, CA. Included are the year of spawning, origin of returning adults based on the presence or absence of a hatchery clip (LM, RM, AD or NO i.e. no clip), the number of individuals successfully genotyped (males and females), the number of individuals excluded [1; missing genotypes at 10 or more loci (total = 201) 2; duplicated genotypes due to multiple sampling of the same individuals (total = 148) 3; identified as Chinook with the species ID locus (total = 7)], and the number of individuals spawned (age three females, age three males and age two males) as reported by the hatchery. *In 2004, only a subset of spawners were sampled for genetic purposes. The total number of females spawned was recorded (276) and it is assumed that approximately 276 males were also spawned based on a female to male spawn ratio of 1:1. Because mated pairs were not recorded, hatchery designation of all spawned individuals (LM, RM, AD or NO) is unknown.

			Genotyped		**Excluded**		**Matings**		
Spawn Year	**Hatchery Clip**	**Hatchery Origin**	**n** [♀]	**n** [♂]	**n** [♀]	**n** [♂]	**n** [♀]	**n** [♂]	**n [jack]**
2004*	LM	IGH	90	99	16	12	-	-	0
	RM	TRH	2	4	1	0	-	-	0
	AD	CRH	0	0	0	0	-	-	0
	None	NO	47	43	3	4	-	-	0
	No info	Unk.	14	13	0	0	-	-	0
	Tot.		153	159	20	16	276	276	0
2005	LM	IGH	158	146	25	18	100	98	0
	RM	TRH	1	0	0	0	0	0	0
	AD	CRH	0	1	0	1	0	0	0
	None	NO	67	83	8	19	4	6	0
	No info	Unk.	9	7	5	4	0	0	0
	Mis-ID	Chin.	3	1	3	1	0	0	0
	Tot.		238	238	41	43	104	104	0
2006	LM	IGH	72	58	0	0	77	56	0
	RM	TRH	1	0	0	0	1	0	0
	AD	CRH	4	9	0	0	2	9	0
	NO	NO	3	5	1	1	4	4	0
	Unk.	Unk.	10	3	9	3	1	0	0
	Tot.		90	75	10	4	85	69	0
2007	LM	IGH	93	142	16	11	125	127	0
	RM	TRH	2	0	1	0	2	0	0
	AD	CRH	0	2	0	0	0	1	0

Continued on next page

Table 2.1 – continued from previous page

Spawn Year	Hatchery Clip	Hatchery Origin	Genotyped n [♀]	Genotyped n [♂]	Excluded n [♀]	Excluded n [♂]	Matings n [♀]	Matings n [♂]	Matings n [jack]
	NO	NO	47	36	1	2	5	5	0
	Unk.	Unk.	77	77	11	12	0	1	0
	Mis-ID	Chin.	1	0	1	0	0	0	0
	Tot.		220	257	30	25	132	134	0
2008	LM	IGH	147	147	6	5	145	142	0
	RM	TRH	0	0	0	0	0	0	0
	AD	CRH	0	0	0	0	0	0	0
	NO	NO	2	4	0	0	3	6	0
	Unk.	Unk.	2	7	0	0	0	0	0
	Tot.		151	158	6	5	148	148	0
2009	LM	IGH	12	16	1	1	13	7	7
	RM	TRH	4	2	1	0	4	2	0
	AD	CRH	0	0	0	0	0	0	0
	NO	NO	5	3	0	0	4	3	0
	Unk.	Unk.	0	0	0	0	0	0	0
	Tot.		21	21	2	1	21	12	7
2010	LM	IGH	219	204	31	22	75	56	8
	RM	TRH	0	1	0	0	0	0	0
	AD	CRH	0	0	0	0	0	0	0
	NO	NO	43	44	9	9	5	14	2
	Unk.	Unk.	1	0	0	0	0	0	0
	Mis-ID	Chin.	1	0	1	0	0	0	0
	Tot.		264	249	41	31	80	70	10
2011	LM	IGH	178	317	8	10	53	68	7
	RM	TRH	0	0	0	0	0	0	0
	AD	CRH	0	1	0	0	0	0	0
	NO	NO	3	17	0	0	4	10	1
	Unk.	Unk.	16	20	1	0	0	0	0
	Mis-ID	Chin.	1	0	1	0	0	0	0
	Tot.		198	355	10	10	57	78	8
2012	LM	IGH	173	434	15	42	64	32	59
	RM	TRH	1	2	0	0	0	0	0
	AD	CRH	0	0	0	0	0	0	0
	NO	NO	5	10	0	1	2	6	1
	Unk.	Unk.	1	4	0	0	0	0	0
	Tot.		180	450	15	43	66	38	60
		Total	3476		353				

Modifications to hatchery spawning procedures took place after 2009 when

adult returns were so sparse (< 60 individuals) that hatchery staff was forced to spawn several jacks and TRH strays for lack of available age-three IGH fish. The small return year coupled with the fact that coho salmon in this basin (SONCC ESU) were already listed as threatened under the Endangered Species Act (1997, updated 2005), prompted the decision to modify the IGH coho salmon program from that of a mitigation program to one focused on coho conservation (HGMP V.10 2013). Subsequently, beginning in 2010, hatchery broodstock were selected based on a genetic spawning matrix designed to reduce inbreeding, jacks were purposely spawned in order to allow gene flow between broodcycles as well as to reduce inbreeding, and hatchery strays from other programs were excluded from matings to encourage local adaptation by ensuring that IGH broodstock represent the natural population with which the program is integrated (HSRG 2012, HGMP V.10 2013). Additionally, natural origin fish were to make up 20-50 % of the broodstock (if/when possible), and males and females were to be spawned at a ratio of 2:1 if fewer than 50 females were available for spawning, in order to increase effective population size (HGMP V.10 2013). These modified hatchery procedures are to be implemented for all future consecutive spawn years in order to protect and conserve the remaining genetic resources in the Upper Klamath River (Chesney and Knechtle 2011, HSRG 2012, HGMP V.10 2013).

2.2.2 Sample Collection and DNA Extraction

A total of 3,476 tissue samples (fin clips) were collected from adult coho salmon returning to IGH from 2004 to 2012. Metadata was collected and recorded for each

sample, and generally included spawn and/or collection date, mate assignment, gender assignment (based on visual identification), fork length, and hatchery mark (Table 2.1). Natural origin fish, or individuals that were not marked by a fin clip were identified in the metadata as NO fish. There were also several fish for which there was no hatchery mark designation or information noted in the metadata, and these were designated as 'Unknown.'

In 2004 only a small fraction (312 of 1734, or 18.0%) of the total return year was sampled for genetic analysis, and spawned individuals were not distinguished as such in the metadata. Beginning in 2005 and lasting until 2009, tissue samples were collected from all spawned individuals as well as a portion of the returning adults not used for broodstock (total fraction of returning adults sampled: 2005, 476 of 1425 (33.4%); 2006, 165 of 332 (50.0%); 2007, 477 of 779 (61.2%); 2008, 309 of 1295 (23.8%); 2009, 42 of 70 (60.0%)). From 2010 through 2012, complete sampling of all broodstock as well as all returning adults not spawned by hatchery staff occurred. Although sampling of the entire hatchery population was not yet standard practice in 2009, due to the extremely low number of returning adults, all LM-clipped fish that returned to the hatchery were spawned and sampled.

DNA from the 3,476 fin samples was extracted using DNeasy 96 tissue kits (QIAGEN Inc.) on a QIAGEN BioRobot 3000. The extractions were then diluted 1:2 in ddH_2O in preparation for a pre-amplification PCR designed to enrich DNA fragments containing the targeted SNPs. The pre-amplification PCR contained the following reagents: 2.5μL of 2X Master Mix (QIAGEN Inc.), 1.3μL of pooled 0.2X unlabeled

primers (derivatives of 96 real-time assays, Applied Biosystems Inc.) and 1.6μL of DNA template. The thermal cycling protocol began with an initial denaturation of 15 min at 95°C, 13 cycles of 15s at 95°C, and 4 min at 60°C for +1°C/cycle. The pre-amplification PCR products were then diluted 1:3 in 2 mM Tris.

2.2.3 Marker Selection and Genotyping

The SNP loci utilized in this study were selected from the 91 SNP assays developed in chapter 1 of this document, as well as from those developed by Smith *et al.* (2006) and Campbell and Narum (2011). The resulting marker panel consists of 95 SNPs, plus 1 species diagnostic assay designed in order to discriminate between coho and Chinook salmon (Table 2.2). Loci in this panel were chosen because their minor allele frequencies (MAFs) in California's Scott Creek coho population were conducive to parentage inference, i.e. MAFs of 0.2 in 80 or more markers (Anderson and Garza 2006). A smaller fraction of these loci were also selected because they demonstrated high pairwise FST values across five coho populations in California, including the Klamath River and IGH, making them useful for population level discrimination (Starks Chapter 1). In the Klamath River, 24 of these SNPs have a minor allele frequency <0.15, 23 between 0.15 and 0.3, and 48>0.3.

Table 2.2: List of the 95 single nucleotide polymorphism loci and 1 coho-Chinook species diagnostic assay used for parentage inference in the 3,476 coho samples collected at Iron Gate Hatchery. Included are the dbSNP accession numbers (at the NCBI on-line repository for short genetic variations) and source reference (SR): 1. Starks Chapter 1; 2. Campbell and Narum 2011; 3. Smith *et al.* 2006.

Locus	**dbSNP**	**S$_R$**	**Locus**	**dbSNP**	**S$_R$**	**Locus**	**dbSNP**	**S$_R$**
Oki94903-192	ss974293192	1	Oki106419-292	ss974293228	1	Oki123470-92	ss974293267	1
Oki95318-100	ss974293193	1	Oki106479-278	ss974293229	1	Oki123921-90	ss974293268	1
Oki96127-66	ss974293194	1	Oki107336-45	ss974293230	1	Oki124162-62	ss974293269	1
Oki96158-278	ss974293195	1	Oki107607-213	ss974293231	1	Oki125998-340	ss974293270	1
Oki96376-63	ss974293197	1	Oki107974-46	ss974293232	1	Oki126160-142	ss974293271	1
Oki97954-228	ss974293199	1	Oki108505-331	ss974293233	1	Oki127236-383	ss974293272	1
Oki100771-83	ss974293200	1	Oki109243-480	ss974293234	1	Oki127760-301	ss974293273	1
Oki100974-293	ss974293202	1	Oki109651-152	ss974293236	1	Oki128302-547	ss974293274	1
Oki101119-1006	ss974293203	1	Oki109874-122	ss974293237	1	Oki128757-232	ss974293276	1
Oki101419-103	ss974293204	1	Oki109894-418	ss974293238	1	Oki128851-185	ss974293277	1
Oki101554-359	ss974293205	1	Oki110064-418	ss974293239	1	Oki129870-552	ss974293278	1
Oki101770-525	ss974293206	1	Oki110078-191	ss974293240	1	Oki130295-48	ss974293279	1
Oki102213-604	ss974293207	1	Oki110689-43	ss974293242	1	Oki130524-184	ss974293280	1
Oki102414-499	ss974293209	1	Oki111681-407	ss974293243	1	Oki131460-243	ss974293282	1
Oki102457-67	ss974293210	1	Oki113457-324	ss974293244	1	Oki131802-368	ss974293283	1
Oki102801-511	ss974293211	1	Oki114315-360	ss974293246	1	Oki_afp4-10	ss263196994	3
Oki102867-667	ss974293212	1	Oki114448-101	ss974293247	1	Oki_arp-105	ss263196997	3
Oki103271-161	ss974293213	1	Oki114587-309	ss974293248	1	Oki_carban-140	ss263197001	3
Oki103577-70	ss974293214	1	Oki116362-411	ss974293249	1	Oki_gdh-189	ss263197003	3
Oki103713-182	ss974293215	1	Oki116865-244	ss974293250	1	Oki_gh-183	ss263197004	3
Oki104515-99	ss974293216	1	Oki117043-374	ss974293251	1	Oki_gshpx-152	ss263197005	3
Oki104519-45	ss974293217	1	Oki117144-64	ss974293252	1	Oki_HGFA-311	ss49845917	2
Oki104569-261	ss974293218	1	Oki117286-291	ss974293253	1	Oki_hsc713-56	ss263197006	3
Oki105105-245	ss974293219	1	Oki117742-259	ss974293254	1	Oki_ins-167	ss49845899	2
Oki105115-49	ss974293220	1	Oki117815-369	ss974293255	1	Oki_itpa-85	ss263197010	3

Continued on next page

Table 2.2 – continued from previous page

Locus	**dbSNP**	$\mathbf{S}_R$	**Locus**	**dbSNP**	$\mathbf{S}_R$	**Locus**	**dbSNP**	$\mathbf{S}_R$
Oki105132-169	ss974293221	1	Oki118152-314	ss974293256	1	Oki_LWSop-554	ss49845907	2
Oki105235-460	ss974293222	1	Oki118175-264	ss974293257	1	Oki_nips-159	ss263197013	3
Oki105385-521	ss974293223	1	Oki118654-330	ss974293258	1	Oki_p53-20	ss263197014	3
Oki105407-161	ss974293224	1	Oki120024-226	ss974293259	1	Oki_pigh-33	ss263197016	3
Oki105897-298	ss974293225	1	Oki120255-113-sppID	ss974293262	1	Oki_rpo2j-235	ss263197019	3
Oki106172-60	ss974293226	1	Oki122593-430	ss974293264	1	Oki_SClkF2R2-120	ss49845937	2
Oki106313-353	ss974293227	1	Oki123205-88	ss974293266	1	Oki_txnip-35	ss263197027	3

All sampled individuals were genotyped with the above SNP panel using the TaqMan chemistry (Applied Biosystems) on a 96.96 Dynamic Genotyping Array with the EP1 Genotyping System (Fluidigm Corporation). Each array included two negative controls (no template), and genotypes were scored using SNP GENOTYPING ANALYSIS SOFTWARE V 3.1.1 (Fluidigm). Additionally, all individuals in this dataset were genotyped with a panel of 16 microsatellite loci (Table 2.3). Microsatellite data was included in this project in order to compare pairwise estimates of relatedness (Rxy) between the two sets of markers. PCRs contained the following reagents in single or multiplex volumes of 15μL with 4μL of 1:20 DNA template: PCR Buffer at 1x concentration 1.5 mM $MgCl_2$, 0.1 mM each dNTP, 0.1 units Amplitaq polymerase (Applied Biosystems), and either 0.17 mM each of forward and reverse primer (single locus PCR) or 0.12 mM of each primer (multiplex PCR). The basic thermal cycling protocol began with a denaturation step at 95°C for 2 min, followed by 10 cycles at 95°C for 15 s, 53°C (the annealing temperature) for 15 s and 72°C for 45 s, then 25 cycles of 89°C for 15 s, 55°C for 15 s and 72°C for 45 s, with a final 5 min extension at 72°C. Modifications to this protocol were made based on the primer pair(s), and involved increasing the annealing temperature to 55°C (followed by 25 cycles at 57°C), decreasing the annealing temperature to 45°C (followed by 25 cycles at 48°C), or simply reducing the 25 cycle step to one with 20 cycles (using the same temperatures as in the basic protocol). The annealing temperature was determined by the primer sequences and prior PCR optimization experiments (Conrad *et al.* 2013). PCR products were amplified and fluorescently labeled for detection on an ABI 3730 DNA Analyzer (Applied Biosystems)

using capillary electrophoresis and standard conditions. For 96 samples from 2005, PCR products were amplified and fluorescently labeled for detection on an ABI 377 Automated DNA Sequencer (Applied Biosystems) using polyacrylamide gel electrophoresis for size separation of the fragments. Allele calling was then carried out by two people independently (in order to check for scoring discrepancies) using Genemapper v. 4.0 (or Genotyper 2.1 for the 96 samples run on the ABI 377; Applied Biosystems). Once genotyping was complete for both marker types, the two datasets were run separately in the Microsatellite Toolkit (Park 2001) in order to identify and remove duplicated samples (i.e. individuals that were sampled more than once at the hatchery).

Table 2.3: Microsatellite loci used to genotype the 3,467 coho samples from IGH in order to compare relatedness values generated from both SNP and microsatellite genotype data. The source references are included

Locus	Reference
OcL 8	Condrey and Bentzen 1998
Oki 1	Smith *et al.* 1998
Oki 13	Smith *et al.* 1998
Omm 1058	Rexroad *et al.* 2002
Omm 1116	Rexroad*et al.* 2002
One μ11b	Scribner *et al.* 1996; R primer redesigned
One μ13	Scribner *et al.* 1996
Ots 103	Beacham and Wood 1999
Ots 1b	Banks *et al.* 1999; R primer redesigned
Ots G3	Williamson *et al.* 2002
Ots G422b	Williamson *et al.* 2002; R primer redesigned
Ots G68	Williamson *et al.* 2002
Ots G83b	Williamson *et al.* 2002
P 53	de Fromentel *et al.* 1992
Ssa 14	McConnell *et al.* 1995, accession no. M75145
Ssa 85	O'Reilly *et al.* 1996

2.2.4 Population Genetics Analyses

Unbiased expected (Hz; Nei 1987) and observed heterozygosity (Ho), as well as the inbreeding coefficient (F_{IS}), were calculated for each return year with both SNP and microsatellite genotypes using Microsatellite Toolkit (Park 2001) and Genetix (Belkhir 2004). Significance of F_{IS} was assessed with 1000 permutations of the dataset.

2.2.5 Parentage Analysis Pedigree Reconstruction

In the Klamath River, following release from IGH, juvenile coho salmon migrate to the ocean to mature and then return to spawn at ages two or three. As such, individuals from brood years 2004-2010 were treated as potential parents of those fish that returned to spawn in 2006-2012. The software package SNPPIT (Anderson 2012)

was used to infer parentage assignments from the individuals in the dataset. This program selects the most likely parent pair for each offspring, then tests for statistical certainty of each parent offspring trio using a novel and efficient (Monte Carlo) simulation method that calculates a p-value and a corresponding false discovery rate (FDR). The genotyping error rate was assumed to be 0.005 per gene copy (1% per locus) for most loci used, however based on Mendelian incompatibilities in reconstructed trios, genotyping error rates could be estimated directly for three loci. Error rates were adjusted for Oki_HGFA-311 to 0.017, Oki106419-292 to 0.007 and Oki129870 to 0.014. Individuals with 10 or more missing loci (85 loci or less) were excluded from the analysis.

For organizational ease, SNPPIT was performed for individual brood years 2006-2012, however each run included all prior years as possible parents. For example, adults that returned in 2009 were treated as potential offspring of those fish that returned in 2004, 2005, 2006, 2007 and 2008. Offspring were not expected to return to the hatchery at age one (offspring from 2009 with putative parents from 2008) nor at age five (offspring from 2009 with parents from 2004), however hypothetical parental years were included in the analyses in order to test for false positive assignments.

Three rounds of SNPPIT analyses were performed. The first was constrained by gender ID and spawn date so that only putative mates of opposite gender and identical spawn date were considered. A second round was then performed which excluded gender ID and spawn date information so that offspring could be assigned to parents of the same sex and/or to parents that spawned on different days. These two runs were then compared in order to identify and resolve metadata errors as well as to confirm correct

assignments. In instances where gender ID disagreements occurred between putative parent-pairs (female-female or male-male), the gender for each individual in question was resolved genetically using the male specific growth hormone pseudogene (GHpsi) marker (Du *et al.* 1993) and PCR. The PCR products were then electrophoresed and visualized on 2% agarose gels. Once the metadata was corrected, a third and final set of constrained SNPPIT runs was then performed for assignment support. An FDR of 0.01 or less was selected as a significance threshold, meaning that no more than 1 of every 100 assignments was expected to be incorrect. Parentage assignments were compared to the mated pairs recorded in the spawning metadata for years 2005-2010.

2.2.6 Age Structure, Reproductive Success and Length of Returning Adults

The age of returning adults was determined for brood years 2007-2012. Individuals that were born in 2004 (the 2004 cohort) were identified when they returned to the hatchery at age three in 2007. Age two fish were not identified in the 2004 cohort because the fraction of broodstock sampled was too small to find the parent-pairs of offspring that returned in 2006. Individuals from the 2005 cohort were identified when they returned at age two or three in 2007 and 2008; individuals from the 2006 cohort were identified when they returned at age two or three in 2008 and 2009; individuals from the 2007 cohort were identified when they returned at age two or three in 2009 and 2010; individuals from the 2008 cohort were identified when they returned at age two or three in 2010 and 2011; individuals from the 2009 cohort were identified when

they returned at age two or three in 2011 and 2012; and individuals from the 2010 cohort were identified when they returned at age two in 2012. The proportion of age two and age three fish from the 2007, 2008 and 2009 cohorts was compared using z-tests. Cohorts 2004, 2005 and 2006 were excluded from this analysis because jacks that were born in these years were purposely not spawned when they returned to the hatchery, so the proportion of age two and age three offspring making up these cohorts would not be a representative sample. Cohort 2010 was also excluded because the age three individuals that returned in 2013 had not yet been analyzed. The parentage analysis was also used to examine the age structure of adult fish that returned to the hatchery in 2010, 2011 and 2012. The number of age two and age three returning adults were counted per year, and z-tests were used to compare the relative proportions of the two age classes.

The distribution of family size was ascertained from the resulting parent-offspring trios for fish that spawned during the period 2004-2010. These calculations included only successful parent pairs that had at least one offspring recovered in the reconstructed pedigrees. Likewise, relative reproductive success was estimated as the number of adult offspring per parent that returned to the hatchery over all years combined. Because the number of offspring per parent pair was not normally distributed, a Kruskal-Wallis test was used to detect differences between the 2007, 2008 and 2009 cohorts as well as between male and female parents within years. Again, only these three cohorts were compared since it is presumed that all putative offspring (two and three year olds) returning to the hatchery were included in the SNPPIT runs corresponding to

these cohorts. The relative reproductive success of males versus females versus parent pairs across all years was also examined as IGH generally reports one-to-one matings (only one male crossed with one female) for the parental years concerned (2004-2010).

Fork lengths were recorded for all sampled salmon returning to the hatchery. This size distribution allowed for comparisons in size ranges based on gender (males versus females) and based on age (age two males versus age three males and females). The size distribution and the parent-offspring trios confirmed in the pedigrees allowed the heritability of length-at-spawning in the dominant three-year old age class to be explored since the parentage analysis identified families. Heritability (h^2) was estimated as the slope of the midparent-offspring regression line. The relationship between the mean fork length of each parent pair and the mean length of all offspring (mean offspring length was calculated for parents with more than one offspring) was compared, as well as all offspring separated by sex. Additionally, the relationships between individual parents and their offspring were examined by specifically looking at the correlations between mothers and all offspring and fathers and all offspring, as well as mothers and daughters and fathers and sons, mothers and sons and fathers and daughters. The relationship between female fork length and relative offspring survival was also explored, since it is generally assumed that larger females produce more eggs, thus potentially more offspring.

2.2.7 Explorations in Relatedness: SNPs versus Microsatellites

Using both SNP and microsatellite genotypes for comparison, Queller and Goodnight's (1989) relatedness coefficient estimator, Rxy, was calculated between all pairs of individuals in each collection year (2004-2012) using the program KINGROUP (Konovalav *et al.* 2004). The Rxy statistic is a prediction of kin relationship, in that it is an estimate of the probability that the alleles shared between two individuals are identical by descent. In theory, the larger the value of Rxy, the greater the extent of relatedness between two individuals. Half siblings have an expected Rxy value of 0.25 and full siblings a value of 0.50. The maximum value of Rxy that can be achieved is 1.0, which indicates identical genotypes.

For each collection year, distributions of Rxy values were plotted, and the mean, standard deviation and skew calculated. Additionally, normal distributions were generated using the means and standard deviations observed in each year, and were compared against the distributions of observed Rxy values. Distributions of Rxy values for all successful parent pairs (pairs with at least one returning offspring assigned) per year were also plotted, and the mean, standard deviation and skew were calculated. Because the hatchery recorded mated pairs from 2005 to 2010, Rxy values for unsuccessful parent pairs (pairs that did not have recovered offspring in the parentage analyses) could be identified. However, distributions were not plotted based on several factors: for cohorts 2005, 2006 and 2010, the total number of 'unsuccessful' parent pairs are not accurate quantifications because the full cohorts were likely not recovered in the

parentage assignments (fractional sampling of spawners in 2005 and 2006, and three-year old offspring that returned in 2013 have not been analyzed); in 2007, although the hatchery reported 266 individuals as being spawned, actual spawn pair information was only available for 152 of these individuals, making it impossible to accurately distinguish unsuccessful pairs from the rest of the un-spawned fish; and for cohorts 2008 and 2009 there were too few unsuccessful pairs to plot distributions of Rxy per year (ten in 2008 and one in 2009). As the Rxy values for successful parent pairs appeared to be normally distributed, a two-sided t-test was used to examine whether the mean values per year and over all years were significantly different between SNPs and microsatellites. The correlation between full sibling family size and Rxy value between parent pairs was also examined separately in both the SNP and microsatellite data.

The parentage analysis identified a prevalence of families consisting of two or more offspring. This allowed for the comparison of pairwise Rxy values between full and half siblings generated with SNP and microsatellite genotype data. Through parentage, it was also observed that a large proportion of the half siblings were sired from one father, an age two male, loosely termed the 'super-jack,' with two females (28 offspring from female one, 41 offspring from female two). As a result, distributions of Rxy were plotted and compared for full and half siblings with and without this male's contribution (all 69 offspring removed) and, as before, the mean, standard deviation and skew were calculated from the observed values and compared against normal distributions produced from the corresponding means and standard deviations. The proportion of type II errors (siblings classified as unrelated) using the two marker types was quantified

using cutoff values of Rxy equal to 0.25 and 0.125 for both full and half siblings. That is, all pairwise values that fell below these two points were summed and divided by the total number of pairwise values in each sibling group. Comparisons of error proportions were then made between the two marker types both with and without the super-jacks's contribution. The frequency of type II errors was then plotted for the entire set of full and half siblings both separately and combined, using both SNP and microsatellite generated values with and without the super-jacks's contribution.

As family size tended to be large overall, simulations of full and half siblings were also performed with only two siblings per parent pair, in order to compare their Rxy values to those generated from the real data. Ten groups of known full and half siblings consisting of 500 individuals (2 siblings per family) were simulated from the allele frequencies obtained from the pooled data set (2004-2012) with the program Nookie (https://github.com/eriqande/nookie). This program uses population allele frequencies to generate male and female genotypes from which full and half sibling offspring can be simulated through Mendelian segregation. The distribution of the Rxy values for the combined group of simulated full and half siblings was plotted, and the mean, standard deviation and skew were calculated.

Additionally, the correlation between Rxy values generated by the two marker types was examined with a simple linear regression for all collection years combined. A similar correlation between Rxy values generated from the two marker types was explored for all siblings, all full siblings and all half siblings (with and without the super-jacks's contribution) with a linear regression. The difference in values of Rxy

generated by the two marker types, or Delta Rxy, was also explored in the full and half siblings. Delta Rxy was calculated by subtracting the SNP generated Rxy values from the corresponding microsatellite values, and the resulting distributions of delta Rxy were plotted for the two sibling groups, both with and without the super-jack's contribution, and the mean, standard deviation and skew were calculated.

2.3 Results

A total of 3,476 coho samples from IGH were extracted and genotyped with the panel of 95 SNPs and one coho-Chinook species diagnostic assay. Of this total, 353 individuals were excluded prior to analysis due to one of three factors: an individual was sampled more than once during the spawn year giving rise to duplicate genotypes (n=145); an individual displayed the Chinook genotype in the species diagnostic assay (n=7); an individual had genotypes that failed at 10 or more loci (n=201, Table 2.1). Coho that strayed from their hatchery of origin to IGH are noted in Table 2.1. This left a total of 3,124 individuals utilized in further analysis. Also, the genotypes produced from Oki109651-152 and Oki110689-43 appeared to be out of Hardy-Weinberg Equilibrium when genotyped on the same panel, so these two loci were excluded from further analyses. The same 352 samples were removed from the microsatellite genotypes, as well as an additional 97 individuals that failed at eight or more loci, leaving a total of 3,027 microsatellite genotypes utilized in further analyses.

2.3.1 Population Genetics Statistics

Using SNPs, estimates of unbiased heterozygosity, Hz, ranged from 0.298 in the 2008 samples to 0.318 in the 2006 samples, and averaged 0.315, while observed heterozygosity, Ho, ranged from 0.301 in the 2005 samples to 0.319 in the 2012 samples and averaged 0.309 (Table 2.4). The inbreeding coefficient, F_{IS}, ranged from -0.0236 in the 2008 sample to 0.0171 in the 2007 sample, and averaged 0.0017 over all years; values were significantly different from zero for the 2004 samples ($P < 0.05$) and the 2007 samples ($P < 0.01$), as well as for all samples combined ($P < 0.001$). The overall degree of relatedness (mean individual Rxy), was estimated by calculating the mean value of Rxy between each individual and all other individuals, and then taking the mean of these individual values in each collection year. The mean individual relatedness ranged from 0.0110 in 2007 to 0.0810 in 2008, and averaged 0.0336 over all years.

For the microsatellite data, estimates of Hz ranged from 0.558 in 2009 to 0.583 in 2004, and averaged 0.569, while Ho ranged from 0.546 in 2011 to 0.576 in 2010, and averaged 0.559 over all years. F_{IS} ranged from -0.0026 in 2010 to 0.0442 in 2011 and averaged 0.0181 over all years; values were significantly different from zero in 2004 and 2006 ($P < 0.01$) and in 2005, 2007, 2008 and 2011, as well as over all years ($P < 0.001$). Mean individual Rxy ranged from -0.0033 in 2004 to 0.0459 in 2009, and averaged 0.0262 over all years (Table 2.4).

Table 2.4: Population genetic statistics of unbiased heterozygosity (Hz), observed heterozygosity (Ho), the inbreeding coefficient (F_{IS}; values with an asterisk are significantly different from zero, 1000 permutations) and the mean individual relatedness coefficient (Rxy; see text) were calculated for each genotyped broodstock year with both SNP and microsatellite markers. Note that * $P < 0.05$, ** $P < 0.01$ and *** $p < 0.001$.

Spawn Year	**n [SNPs]**	**Hz**	**Ho**	**F_{IS}**	**Mean indiv. Rxy**	**n [Msats]**	**Hz**	**Ho**	**F_{IS}**	**Mean indiv. Rxy**
2004	276	0.312	0.308	0.0115*	0.0220	273	0.583	0.572	0.0188**	-0.0033
2005	392	0.302	0.301	0.0040	0.0597	377	0.565	0.553	0.0222***	0.0368
2006	151	0.318	0.315	0.0089	0.0291	150	0.564	0.552	0.0224**	0.0458
2007	422	0.314	0.309	0.0171**	0.0110	401	0.578	0.567	0.0180***	0.0074
2008	298	0.298	0.305	-0.0236	0.0810	285	0.565	0.552	0.0233***	0.0453
2009	39	0.310	0.306	0.0147	0.0217	37	0.558	0.555	0.0058	0.0459
2010	441	0.315	0.314	0.0054	0.0124	434	0.575	0.576	-0.0026	0.0117
2011	533	0.305	0.306	-0.0038	0.0456	507	0.572	0.546	0.0442***	0.0226
2012	572	0.313	0.319	-0.0187	0.0196	563	0.567	0.561	0.0108	0.0233
	mean	0.310	0.309	0.0017	0.0336	**mean**	0.569	0.559	0.0181	0.0262

2.3.2 Pedigree Reconstruction

The initial round of SNPPIT runs was constrained by spawn date and gender ID and generated pedigrees comprised of 1,470 putative father-mother-offspring trios. An FDR of 1 in 100 (0.01) was chosen as a basis for excluding false positive trio assignments. Ultimately, trios were removed if they possessed an FDR value greater than or equal to 0.01 with a corresponding low maximum posterior probability.

An unconstrained set of runs was then performed with spawn date and gender ID excluded. This second round produced a total of 1,512 putative father-mother-offspring trios. Of the 42 additional trios resulting from the unconstrained runs (comprised of 15 unique parent pairs), 23 displayed discrepancies in the parent's gender ID, 6 in the parent's spawn date, and 13 in both the parent's gender ID and spawn date. All of these trios had FDR values ranging from 0 to 0.014 and maximum posterior probabilities ranging from 0.753 to 0.999. As expected, in both the constrained and unconstrained analyses, returning adults were never inferred as being less than two years of age nor greater than three years of age.

Discrepancies in gender ID were resolved using the male specific growth hormone pseudogene (GHpsi) marker (Du *et al.* 1993). PCR products were electrophoresed and visualized on 2% agarose gels. Of the 36 additional trio assignments with discrepancies in gender ID, only one trio appeared to be an incorrect assignment. The offspring individual in this trio assigned to a male-male parent pair in the unconstrained analysis and was excluded from the constrained analysis. When electrophoresed with the GH-

psi marker, both parental individuals in the male-male trio displayed the male band. According to the hatchery, both males in this trio were spawned with two different females, however a genetic sample was never obtained for one of the females. As such, it is likely that one of the paternal males in this mis-assigned trio was a close relative of the true mother due to the low FDR (0.0008) and p value (0.001), and the high maximum posterior probability (0.985) associated with this false assignment. Ultimately, this male-male trio was removed from the pedigree results.

In the remaining trios with spawn date discrepancies, it was assumed that note-taking errors in the meta-data occurred. These errors were corrected appropriately when possible. All other trios generated from the unconstrained runs corresponded to the gender and spawn date metadata associated with the parent pairs in the original constrained analysis. One final round of pedigree analysis was then performed, and was constrained by the newly resolved metadata.

The final number of parent-offspring trios identified in the pedigree analysis was 1,511. FDR scores ranged from 0 to 0.0078 (mean, 6.38×10^{-5}), with p-values ranging from 0 to 0.05 (mean, $8.00\text{x}10\times10^{-4}$) and posterior probabilities ranging from 0.5454 to 0.9999 (mean, 0.9896). In this analysis an FDR of 0.0078 means that no more than 12 of the 1,511 parentage assignments are expected to be incorrect, although there were only 8 assignments with FDR scores greater than 0.002. Furthermore, there were no instances in which offspring were falsely assigned to parents as juveniles (less than two years of age) or as adults greater than four years of age (only two- and three-year-olds recovered) in either the constrained or unconstrained pedigree reconstruction, providing

further support for this high confidence. Additionally, because the hatchery recorded mated pairs (for most of the spawned individuals), I was able to test for mismatches in the hatchery provided spawning information with the actual spawners that were recovered in the pedigrees. Of the 381 mated pairs (with spawn info) 29 (7.6%) were recovered that were not recorded as being crossed in the hatchery data.

The 1,511 reconstructed parent-offspring trios correspond to 72.2% of the total adult offspring from 2007 to 2012 assigned to a parent-pair from 2004 to 2010 (Table 2.5). This is because only a small fraction of the 2004 broodstock was sampled and available for analysis. When the 2007 return year is excluded from the total, the remaining 1,455 parent-offspring trios represent 83.0% of adult offspring assigned to a parent pair. The remaining 299 offspring were likely not assigned parentage for one of several reasons: they were either of natural origin (NO) or they were strays from the Trinity River (RM) or Cole M Rivers hatcheries (AD); the true parents of these individuals were excluded prior to analysis due to poor data quality (missing 10 $>$ loci); or the parent's genetic sample was not received (2 individuals; Table 2.1). Because offspring were assigned to parent pairs, and not to single parents, if one spawner was excluded from the analysis, that spawner's mate was excluded as well and the opportunity for accurate offspring assignment was no longer possible for the pair. As the hatchery kept records of the matings from 2005 to 2010, spawned individuals that were excluded from the analysis could be determined and, from those, the number of offspring that would have been assigned had these individuals been available could be estimated (Table 2.5). This was achieved by first calculating and subtracting the number of excluded parent pairs

(excluded males x excluded females) per spawn day, from the calculated number of genotyped parent pairs (all females x all males), and then weighing the value by the proportion of females spawned on that day. When these values were summed over the spawn year, the percentage of parent pairs available for analysis was estimated, and then used to scale observed offspring recoveries. The scaled estimates of offspring recoveries from 2008 to 2012 indicate that had all parent pairs been included in the analysis, an additional 165 offspring, or a total of 1,620 individuals, would have been assigned parentage, which would be 92.4% of all 2008-2012 offspring in the dataset.

Table 2.5: Summary of offspring (offs.) genetically assigned to parents from seven spawn years of coho salmon at Iron Gate Hatchery. Listed are the number of females and males (age-two and -three) with recovered offspring, the proportion of parent-pairs included in the sample collection (see text), the number of mated pairs recorded at the hatchery and the number of those matings confirmed by pedigree analysis. The number of offspring recovered by parent pairs is expected to be higher in more recent years because, beginning in 2009, I analyzed all IGH returns, not just the spawners. A scaled estimate of the number of expected recoveries had all spawners been included in the analysis is also shown, *In 2004 mated pairs were not designated in the spawn data, reported are the pairs that were assigned offspring through the parentage analysis.

Spawn Year	**Females with Offs.**	**Males with Offs.**	**Jacks with Offs.**	**% Parent Pairs Included**	**Rec. Mates [hat.]**	**Conf. Mates [PBT]**	**Offs. recovered [# analyzed] in year**						**Total [2093]**
							2007 [339]	**2008 [292]**	**2009 [26]**	**2010 [370]**	**2011 [513]**	**2012 [553]**	
2004*	18	18	0	0.065	276	20	24	0	0	0	0	0	24
2005	84	82	0	0.930	104	89	32	245	0	0	0	0	277
2006	16	17	0	0.824	84	17	0	2	17	0	0	0	19
2007	76	78	0	0.841	134	79	0	0	5	249	0	0	254
2008	111	109	0	0.923	148	115	0	0	0	41	341	0	382
2009	16	11	5	0.857	20	18	0	0	0	0	84	170	254
2010	59	51	9	0.952	80	68	0	0	0	0	0	301	301
sum	380	366	14	**sum**	846	406	56	247	22	290	425	471	1511
						scaled	-	266	27	341	468	518	

The reconstructed trios made it possible to identify one four-generation family with two offspring, two parents, two maternal grandparents and two paternal great-grandparents. A total of 90 three-generation families were identified in which offspring (mean = 4.04, range = 1-25 per family), parents and the four grandparents were known, while a total of 54 families were identified in which the offspring (mean = 4.61, range = 1-29 per family), parents and only two grandparents were known. Of the latter there were 22 families in which the paternal grandparents were identified and 32 in which the maternal grandparents were identified.

Direct assessment of certain hatchery practices was also possible with the parentage analysis. The reconstructed pedigrees confirmed one instance of inbreeding, in which one full sibling parent-offspring trio was recovered. Although spawning hatchery strays has recently been terminated (HGMP V.10 2013), the analysis identified a total of 7 RM clipped fish and 1 AD clipped fish spawned prior to 2010. In all except one instance (in 2009, 1 RM by RM cross), hatchery strays were crossed with LM clipped fish. Incorporating jacks into the broodstock has also been avoided at IGH prior to 2010 (HSRG 2012, HGMP V.10 2013), however in 2009 due to the small return year, 5 of the 16 (31.3%) successful fathers recovered in the pedigree were precocious two-year-olds. In 2010, 9 of the 60 (15.0%) successful fathers recovered were jacks. A total of 43 NO spawners (18 of which were spawned in 2010 alone) were also identified in the pedigrees. In all except one mating (2009, RM by NO cross), NO fish were crossed with LM fish. Lastly, the analysis identified hatchery mark ambiguities. In three instances, samples did not receive a hatchery clip (NO) and were recovered in the pedigrees as IGH origin

fish. In two instances, fish were recorded as receiving both an RM and an LM clip and were recovered in the pedigrees as IGH origin fish. Additionally, there were 39 samples for which there was no hatchery clip information provided in the metadata (1 sample labeled "K") which were recovered in the pedigrees as IGH origin fish.

2.3.3 Age Structure

The age composition of adults returning to spawn was evaluated for individuals conceived in 2007, 2008 and 2009. Of the 253 fish assigned to parents spawned in 2007 (cohort 2007), a total of 5 (1.98%) returned as jacks and 248 (98.02%) returned as three-year-olds (43.1% males and 55.3% females). Of the 382 fish that assigned to parents spawned in 2008 (cohort 2008), a total of 41 (10.7%) returned as jacks and 341 (89.3%) returned as three-year-olds (48.2% males and 41.1% females). Of the 254 fish that assigned to parents spawned in 2009 (cohort 2009), 84 (33.1%) returned as jacks and 170 (66.9%) returned three-year-olds (21.3% males and 45.6% females; Figure 2.2). For all three cohorts, there were significant differences in the proportion of male fish that returned at age two ($z_{2007:2008}$ = -4.17, $P < 0.001$; $z_{2008:2009}$ = -6.94, $P < 0.001$; $z_{2007:2009}$ = -9.20, $P < 0.001$). Although age three adults composed the majority of all three cohorts, significantly more males from the 2009 cohort returned to spawn at age two than at age three ($z = 3.80$, $P < 0.001$). There were also significant differences in the proportion of age three males in cohorts 2007 and 2009 ($z = 5.39$, $P < 0.001$) and cohorts 2008 and 2009 ($z = 7.01$, $P < 0.01$), as well as a significant difference in the proportion of females that returned at age three for cohorts 2007 and 2008 ($z = 3.69$,

$P < 0.001$). Additionally, there were significantly more age three females than both age two and age three males combined in the 2007 cohort ($z = 2.22$, $P < 0.05$) while in the 2008 cohort there was a significantly greater proportion of males than females ($z = -4.92$, $P < 0.001$). There were more males than females in the 2009 cohort, however this difference was not significant. Two-year old females were not recovered in either of the three cohorts, nor in the rest of the analysis.

The age composition of adults returning to the hatchery to spawn in 2010, 2011 and 2012 was also examined. Each return year consisted of age two and age three adults that were conceived in two separate years: 2007 or 2008 (2010), 2008 or 2009 (2011) and 2009 or 2010 (2012). Of the 290 fish that returned to the hatchery in 2010, 41 (14.1%) were jacks and 249 (85.9%) were three-year-olds (43.8% males and 56.2% females). Of the 425 fish that returned to the hatchery in 2011, 84 (19.8%) were jacks and 342 (80.5%) were three-year-olds (54.1% males and 45.9% females). Of the 471 fish that returned in 2012, 301 (63.9%) were jacks and 170 (36.1%) were three-year-olds (31.8% males and 68.2% females; Figure 2.3). Significantly more jacks returned in 2012 than 2010 ($z = 13.4$, $P < 0.001$) and in 2011 ($z = 13.3$, $P < 0.001$). The proportion of jacks that returned in 2011 was greater than in 2010 but the difference was not significant ($z = -1.95$, $P > 0.05$). The proportion of females decreased with each return year, and differed significantly between all three years ($z_{2010:2011} = 3.02$, $P < 0.01$; $z_{2010:2012} = 6.71$, $P < 0.001$; $z_{2011:2012} = 4.00$, $P < 0.001$). Although the proportions of males and females were close to equal in the 2010 return year, there were significantly more age three females than age three males ($z = 2.78$, $P < 0.01$) while in

the 2011 return year there were significantly more age three males than females ($z = 2.14$, $P < 0.05$). In the 2012 return year there were significantly more age three females than age three males ($z = 6.72$, $P < 0.001$), however jacks significantly dominated the 2012 return year ($z = 8.54$, $P < 0.001$).

2.3.4 Family Size and Relative Reproductive Success

The 1,511 parent-offspring trios consisted of 405 parent-pairs distributed in 358 pedigrees and included 379 male parents and 380 female parents. Although spawning was generally performed at a male to female ratio of 1:1 (Hampton 2005; Richey 2006; Chesney 2007; Chesney 2009; Chesney and Knechtle 2011), 23 males were found to have been mated with more than one female (all 1 x 2 crosses except for one 1 x 3 cross) and 24 females were found to have spawned with more than one male (all 1 x 2 crosses). The mean number of offspring for all successful parent pairs was 3.71 (range 1-41), while the mean number of offspring for successful parent pairs in cohorts 2007, 2008 and 2009 was 4.21 (range, 1-41; Figure 2.4). The mean number of offspring for all successful males was 4.00 (range 1-69), while the mean for all successful males in cohorts 2007 to 2009 was 4.54 (range 1-69). The mean number of offspring for all successful females was 3.99 (range 1-41), and the mean for successful females from the 2007 to 2009 cohorts was 4.38 (range 1-41). In all years combined, only 32.5% of all parent pairs had a single offspring return to the hatchery (females: 28.1%, males: 27.7%) and when only considering parents from 2007 to 2009 this proportion decreased to 26.4% (females: 22.7%, males: 22.7%). The male parent with the highest reproductive success

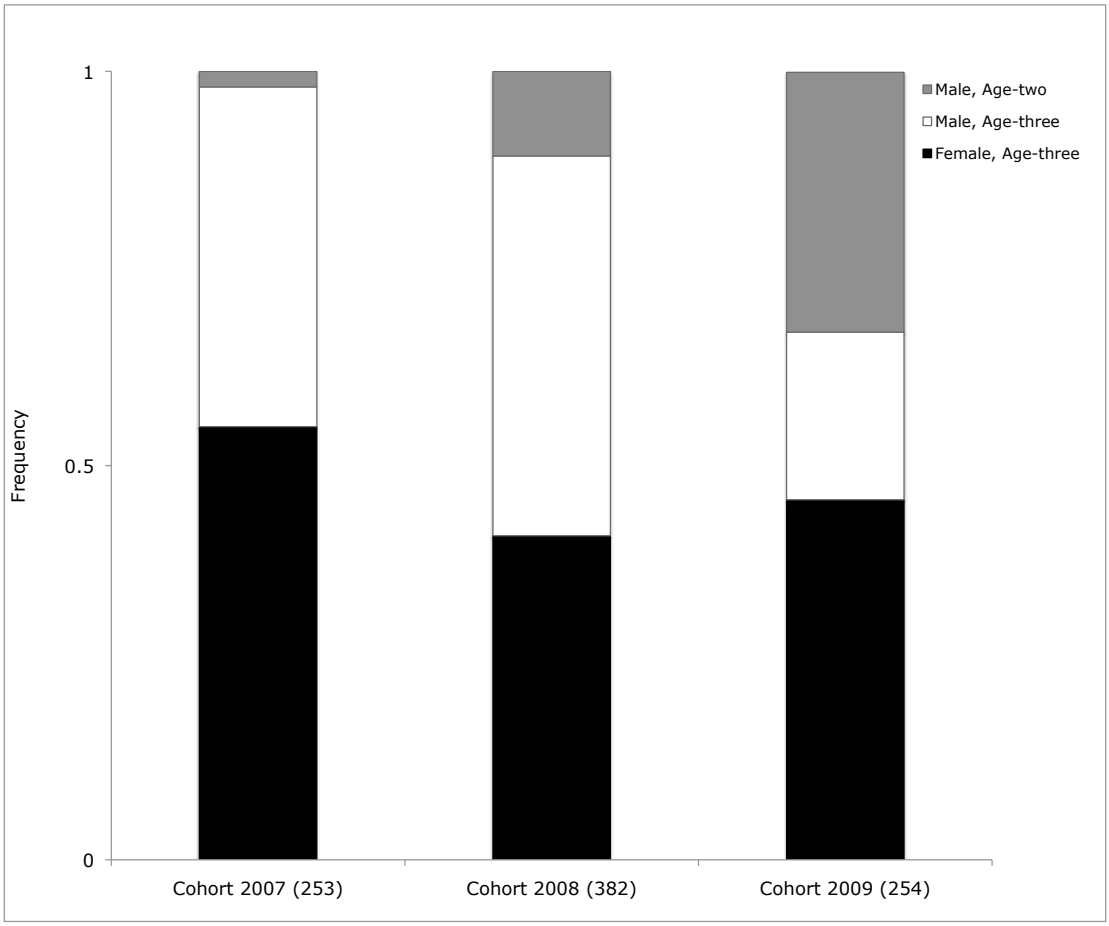

Figure 2.2: Age distribution of returning adults (male and female) for three cohorts (2007, 2008 and 2009) from IGH. Numbers in parentheses indicate the total number of fish. Grey bars represent age-two males, white bars represent age-three males and black bars represent age-three females.

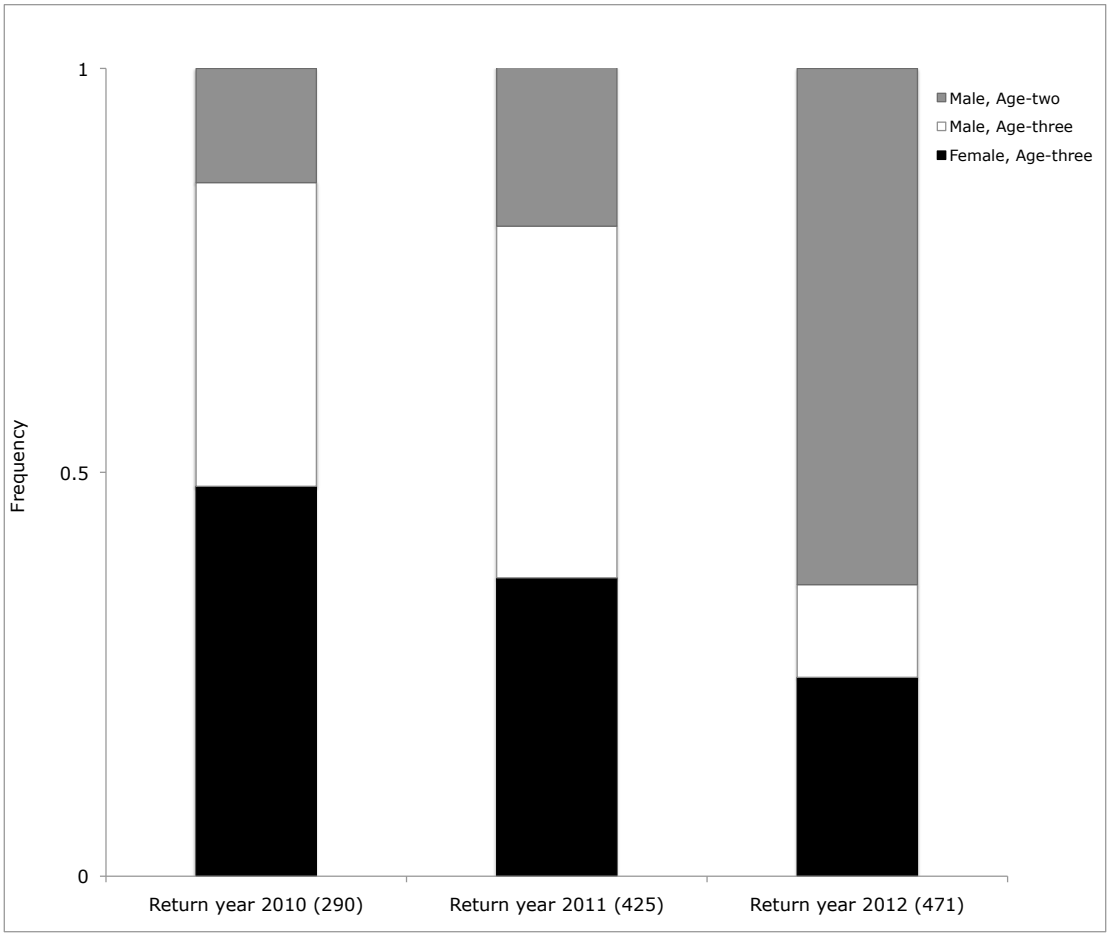

Figure 2.3: Age distribution of adults (male and female) that returned to IGH in 2010, 2011 and 2012. Numbers in parentheses indicate the total number of fish. Grey bars represent age-two males, white bars represent age-three males and black bars represent age-three females.

(69 offspring) was a precocious two-year old (the 'super-jack') with a fork length of 380 mm from the 2009 return year. This individual was mated with two females, one of which had the largest reproductive success for a female (41 offspring) and produced the largest full sibling family recovered in the entire analysis.

In 2004, there were too few spawners sampled to achieve good representation of family size for that year (Table 2.1) and in 2006 only 20.2% of spawned individuals had offspring recovered (probably an underestimation due to fractional sampling in 2008, i.e. likely did not recover all two-year olds). In both of these years, the proportion of parent pairs with only one offspring was much greater than when all years were combined. Additionally, in 2009, due to the small number of spawned individuals, there were more parent pairs with eight returning offspring than with a single offspring. Aside from these three years, the distribution of family sizes across years was comparable to the distribution of reproductive success (Figure 2.5). There was no significant difference between male and female mean reproductive success, and although males had a greater coefficient of variation likely driven by the 'super-jack' and his offspring ($CV_{2007:2009}$: $males_{CV} = 1.38$, $females_{CV} = 1.15$, $F_{1,1405} = 1.45$, $P = 0.956$; $CV_{AllParents}$: $males_{CV} = 1.27$, $females_{CV} = 1.11$, $F_{1,1758} = 1.31$, $P = 0.996$), this difference was not significant. Lastly, the correlation between female length and the number of her offspring that returned did not produce a significant relationship.

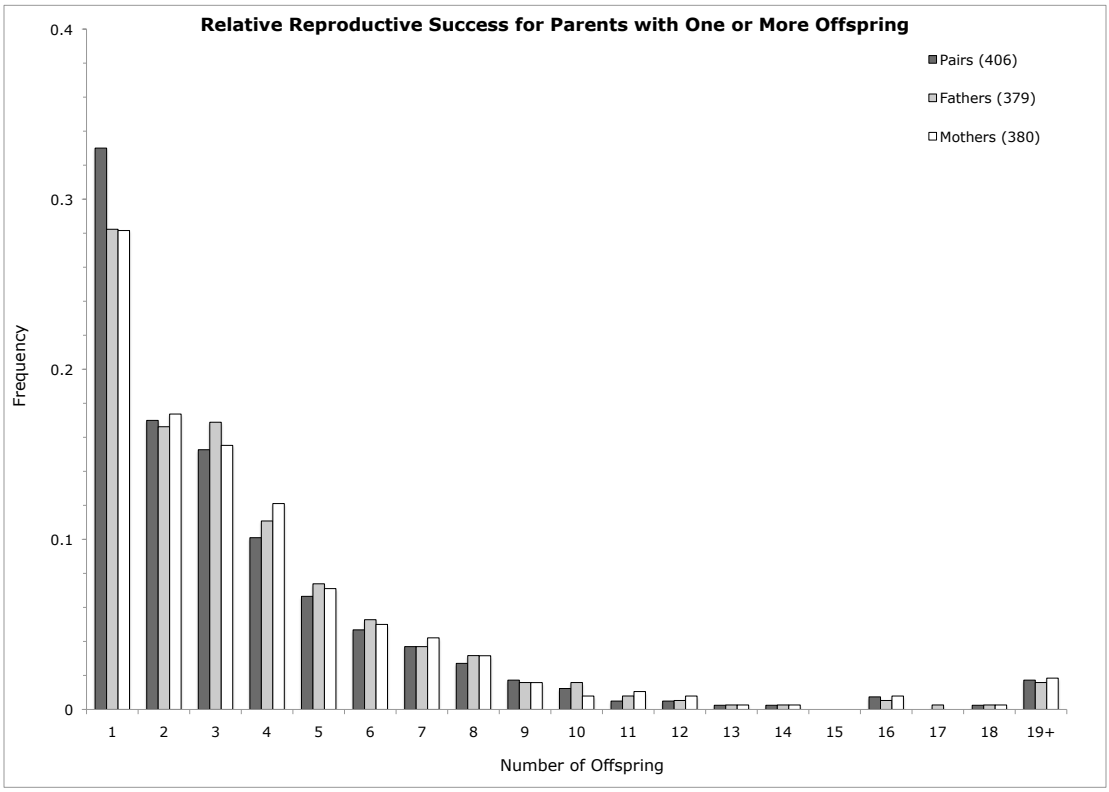

Figure 2.4: The number of offspring that returned to the hatchery for mated pairs (dark grey bars), fathers (light grey bars) and mothers (white bars) over all sample years. The numbers in parentheses indicate the total number of mated pairs, fathers and mothers.

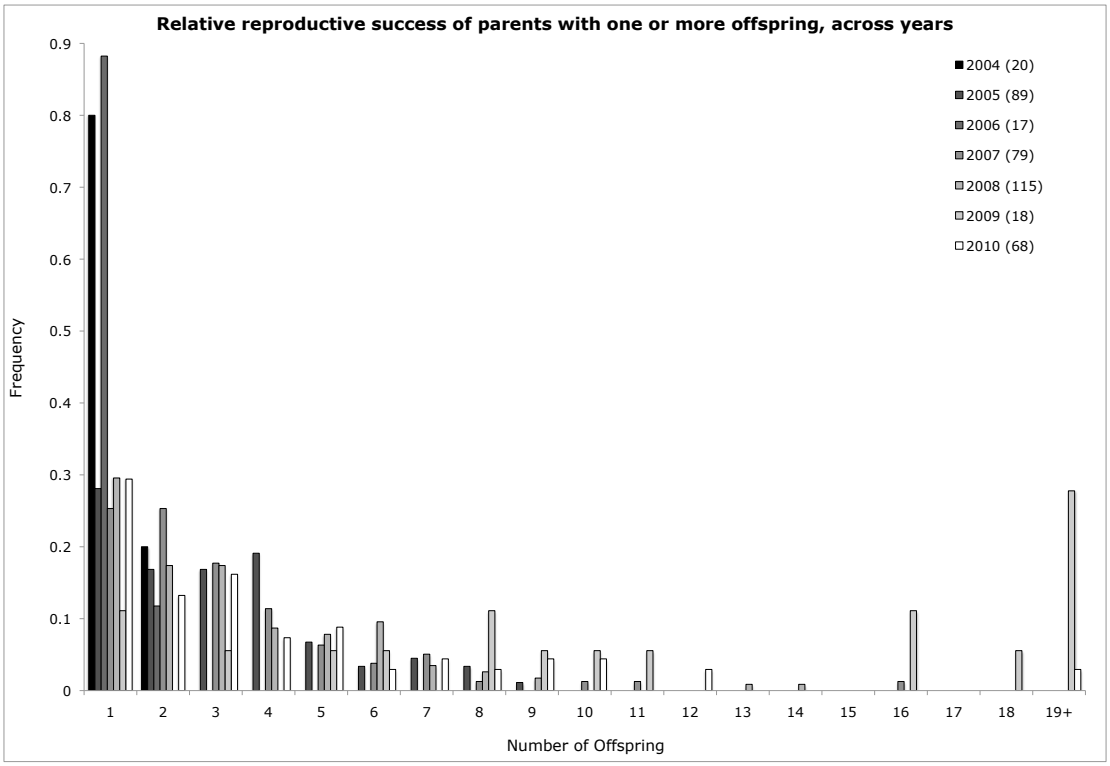

Figure 2.5: The number of full-sibling offspring that returned to the hatchery for parents spawned in each year, 2004-2010. Note that three-year old offspring of the 2010 spawners have not been analyzed, so the offspring assigned to this year are underrepresented.

2.3.5 Distribution of Fork-Length and Heritability of Length-at-spawning

Using the reconstructed pedigrees and the lengths of sampled individuals, the distribution of size by age and gender for all adults was assessed. The mean length for the total 471 age-two males that returned to the hatchery was 438.51 mm (range 340-690mm), the mean length of the total 749 age-three males that returned to the hatchery was 695.21 mm (range 440-810 mm), and the mean length of the total 822 age-three females that returned to the hatchery was 681.85 mm (range 510-790 mm). Despite two outliers (1 age-three male measuring 440 mm and 1 age-two male measuring 690 mm) that likely represent errors in the recorded metadata, there is a clear distinction in size distribution between age-three adults and age-two males (Figure 2.6).

Using the families inferred from the parentage analysis, parent-offspring regressions of length were examined for three-year old offspring in the following comparisons: parental mean - all age-three offspring, parental mean - female offspring, parental mean - age-three male offspring, mother - all offspring, father - all offspring, mother - daughter, father - son, mother - son and father - daughter (Figure 2.7 and Figure 2.8). All comparisons suggested small positive correlations between parental and offspring length, however only three of the nine comparisons were significant ($P < 0.001$), although R^2 was always small. Mean parental length explained approximately 5.4% of the observed variation for all age-three male offspring ($F_{1,230} = 14.29$, $R^2 = 0.054$, $P < 0.001$), however the signal of this pattern decreased when all age three offspring were considered ($F_{1,322} = 3.582$, $R^2 = 0.008$, $P = 0.059$), and it disappeared when considering only female

offspring ($F_{1,239}$ = 0.454, R^2 = -0.002, P = 0.50). When examining parents separately, maternal length explained approximately 8.9% of the observed variation in all offspring ($F_{1,309}$ = 31.22, R2 = 0.089, P < 0.001) and approximately 5.8% of the variation in male offspring length ($F_{1,231}$ = 15.26, R^2 = 0.058, P < 0.001). Heritability (h^2) was calculated as the slope of the parent-offspring regression line. Although heritability of mean parent length in relation to all offspring (h^2 = 0.103), and heritability of paternal length in relation to male offspring (h^2 = 0.104), was relatively large, it was highest for mean parent length and all male offspring (h^2 = 0.298), followed by maternal length and all offspring (h^2 = 0.248) and maternal length and all male offspring (h^2 = 0.237; Table 2.6).

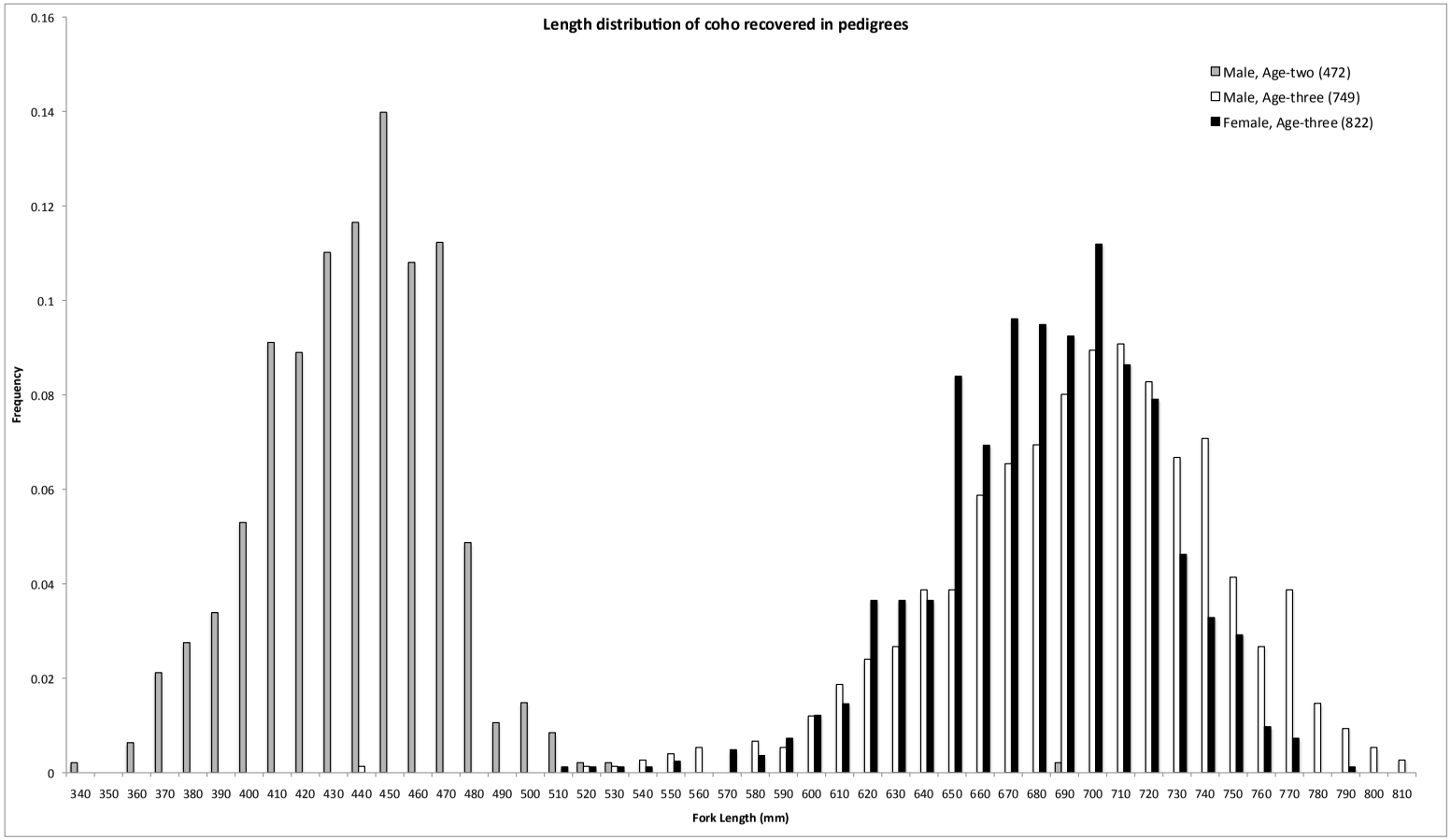

Figure 2.6: Length distribution of all adults recovered with parentage. Grey bars indicate two-year old males, white bars indicate three-year old males and black bars indicate three-year old females. Numbers in parentheses indicate the number of individuals used in the analysis.

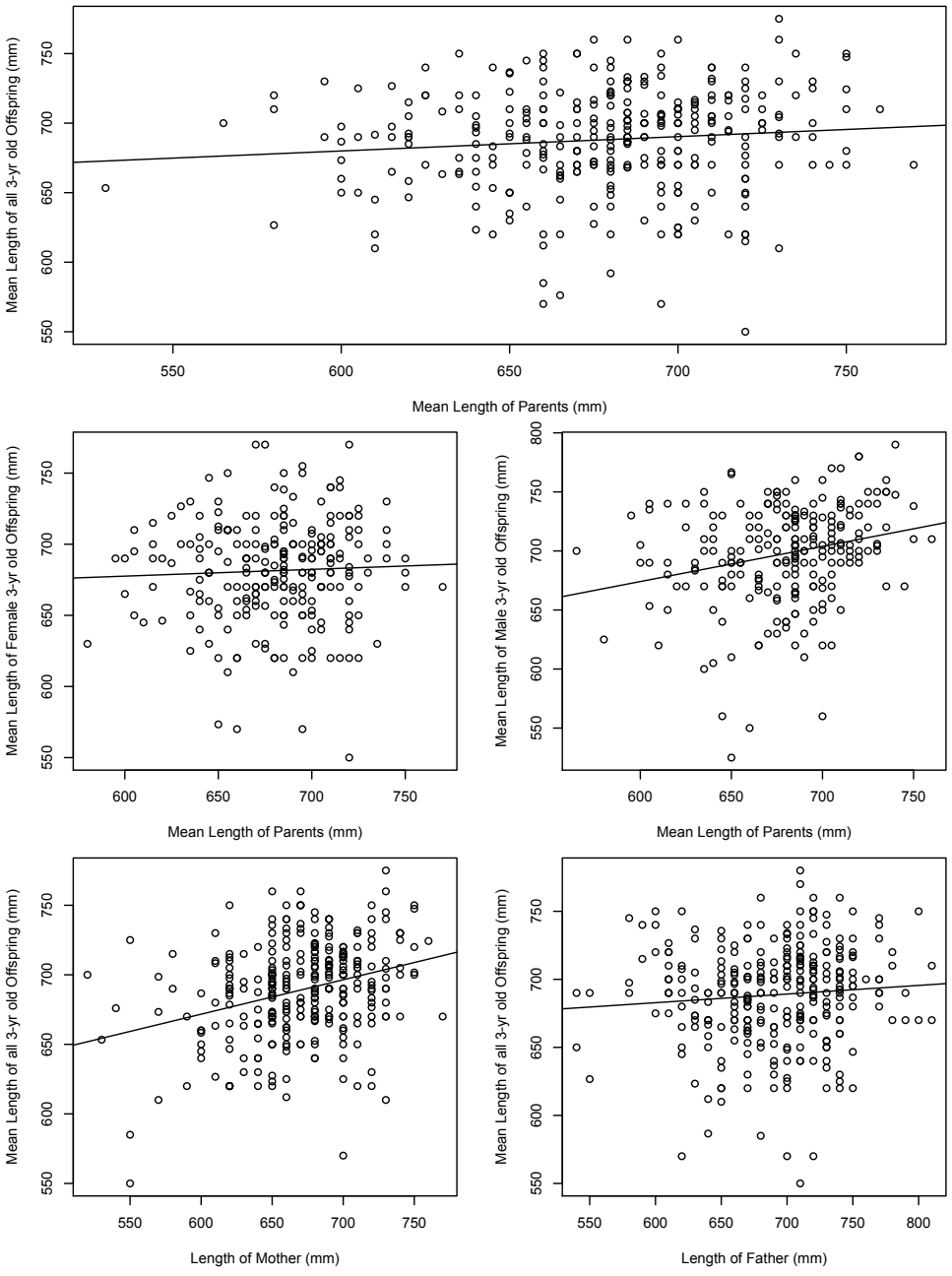

Figure 2.7: Linear regression of parental length on 3-year old offspring length. Independent comparisons were made for: mean parent length and all offspring, female offspring and male offspring, as well as, mothers and all offspring and fathers and all offspring.

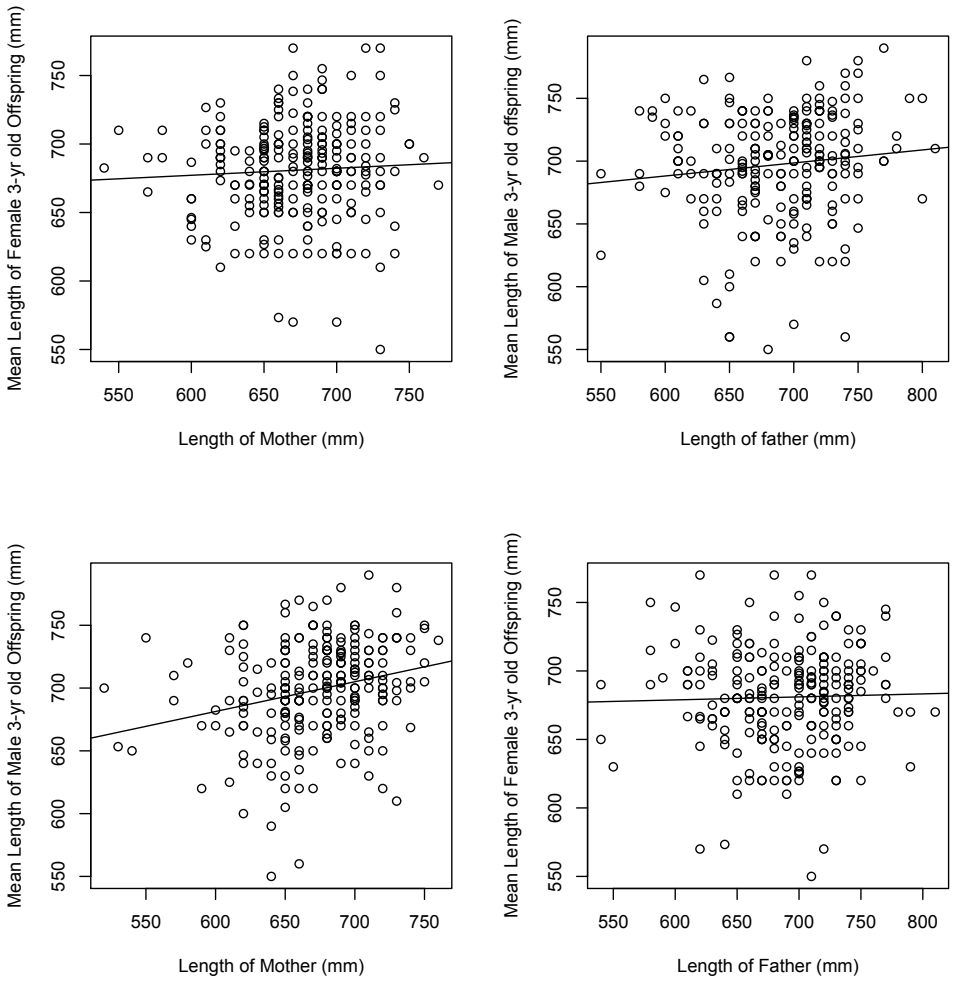

Figure 2.8: Linear regression of parental length on 3-year old offspring length. Independent comparisons were made for: mothers and female offspring, fathers and male offspring, mothers and male offspring and fathers and female offspring.

Table 2.6: Heritability (h^2) of length-at-maturity estimated as the slope of the parent-offspring regression line between different comparisons of parents and offspring (Figure 2.7 and Figure 2.8). Mean is the average length of the parents, and the regression goodness of fit (R^2) and standard error (SE) are also reported. *** indicates significance at $P < 0.001$.

Parent	**Mean**	**Mean**	**Mean**	**Female**	**Male**	**Female**	**Male**	**Male**	**Female**
Offspring	**All**	**Female**	**Male**	**All**	**All**	**Female**	**Male**	**Female**	**Male**
h^2	0.103	0.048	0.298***	0.248***	0.064	0.051	0.104	0.022	0.237***
R^2	0.008	-0.002	0.054	0.089	0.004	-0.001	0.009	-0.003	0.058
SE	36.76	36.02	40.94	33.02	36.51	35.71	42.38	36.16	38.05

2.3.6 Explorations in Relatedness

The Rxy statistic was used to calculate relatedness between all pairs of individuals within each collection year using both SNP and microsatellite genotypes for comparison. For SNPs, over all samples, Rxy ranged from -0.62 to 0.89 (mean = 0.031), while the mean of all pairwise Rxy values within each broodstock collection (2004-2012) ranged from 0.0099 in 2007 to 0.0797 in 2008. Distributions of Rxy over all years and for each collection were normally distributed, although skews were positive, ranging from 0.054 in 2007 to 0.280 in 2008 (mean = 0.223), except in 2012 where the distribution was not normal and was very positively skewed (0.701; Figure 2.9). A positive skew indicates an asymmetry towards Rxy values greater than zero, which can be seen in the longer tail of more positive relatedness estimates. For microsatellites, over all samples, Rxy ranged from -0.60 to 0.90 (mean = 0.020), while the mean of all pairwise Rxy values within each broodstock collection (2004-2012) ranged from -0.0032 in 2004 to 0.0479 in 2006. Distributions of Rxy over all years and for each collection, were not normally distributed and the skew was large, ranging from 0.218 in 2009 to 0.624 in 2012 (mean = 0.396; Figure 2.10).

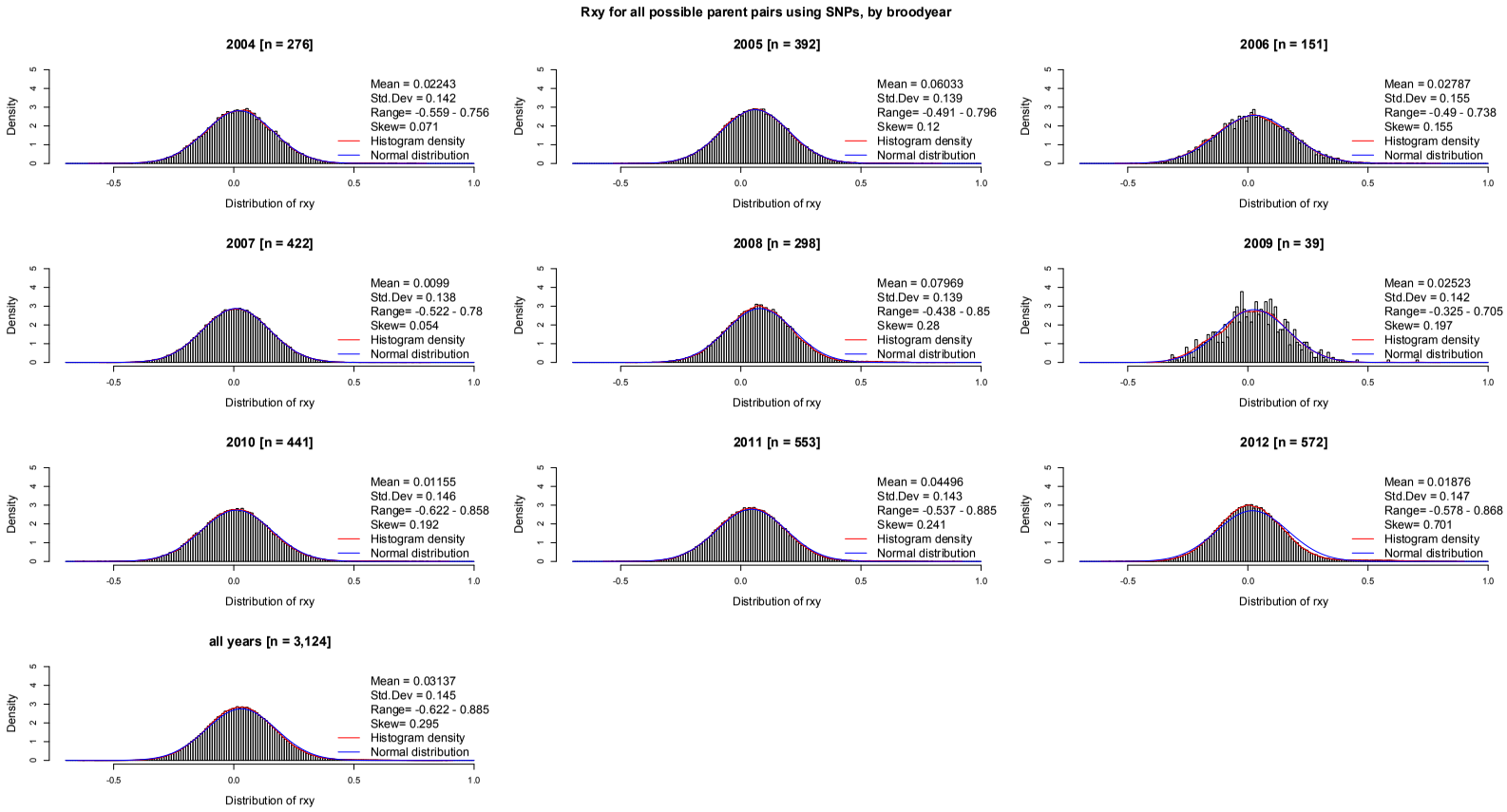

Figure 2.9: Distribution of Rxy, between all possible pairs of individuals in each collection year and over all years generated with SNPs (n is the total number of individuals). The mean, standard deviation (Std. Dev.), range and skew are reported.

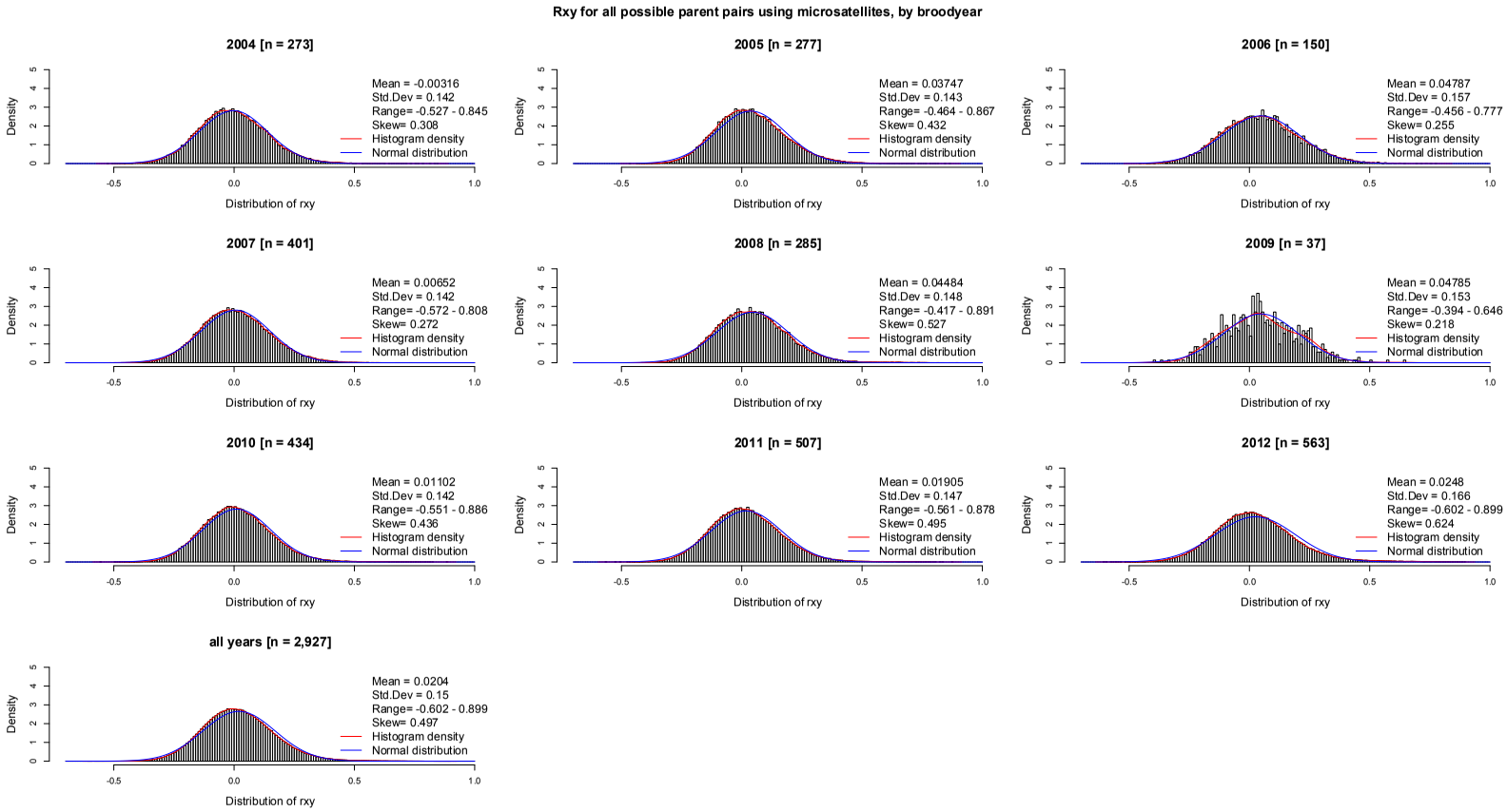

Figure 2.10: Distribution of Rxy between all possible pairs of individuals in each collection year and over all years generated with microsatellites (n is the total number of individuals). The mean, standard deviation (Std. Dev.), range and skew are reported.

The distribution of relatedness values between all parents assigned offspring generated with SNP genotypes was compared to values of relatedness generated with microsatellites for the same individuals. For all successful parent pairs across years, SNP-generated Rxy values ranged from -0.36 to 0.46 (mean = 0.0585) with a skew of -0.052, and from -0.51 to 0.49 (mean = -0.0042) with a skew of 0.250 for microsatellite values. Mean Rxy for all years (2004 - 2010) was greater for SNP (range = -0.0283 - 0.0911) than for microsatellite data (range = -0.0750 - 0.0258), and these differences were significant in 2005, 2009 and 2010 ($P < 0.001$; t-test), and over all years ($P < 0.001$; t-test). With the exception of 2006 and 2009, skew was generally less positive in the SNP-generated values (range = -0.228 - 0.229; Figure 2.11) than in those generated by microsatellites (range = -0.305 - 0.503; Figure 2.12). A weak negative correlation was detected between the degree of relatedness of parent pairs (all years) and the number of offspring they produced that returned to the hatchery for both SNP- ($r = -0.1635$, $P < 0.001$) and microsatellite-generated values of Rxy ($r = -0.1382$, $P < 0.01$), suggesting that less related parents may have greater reproductive success (Figure 2.13 and Figure 2.14).

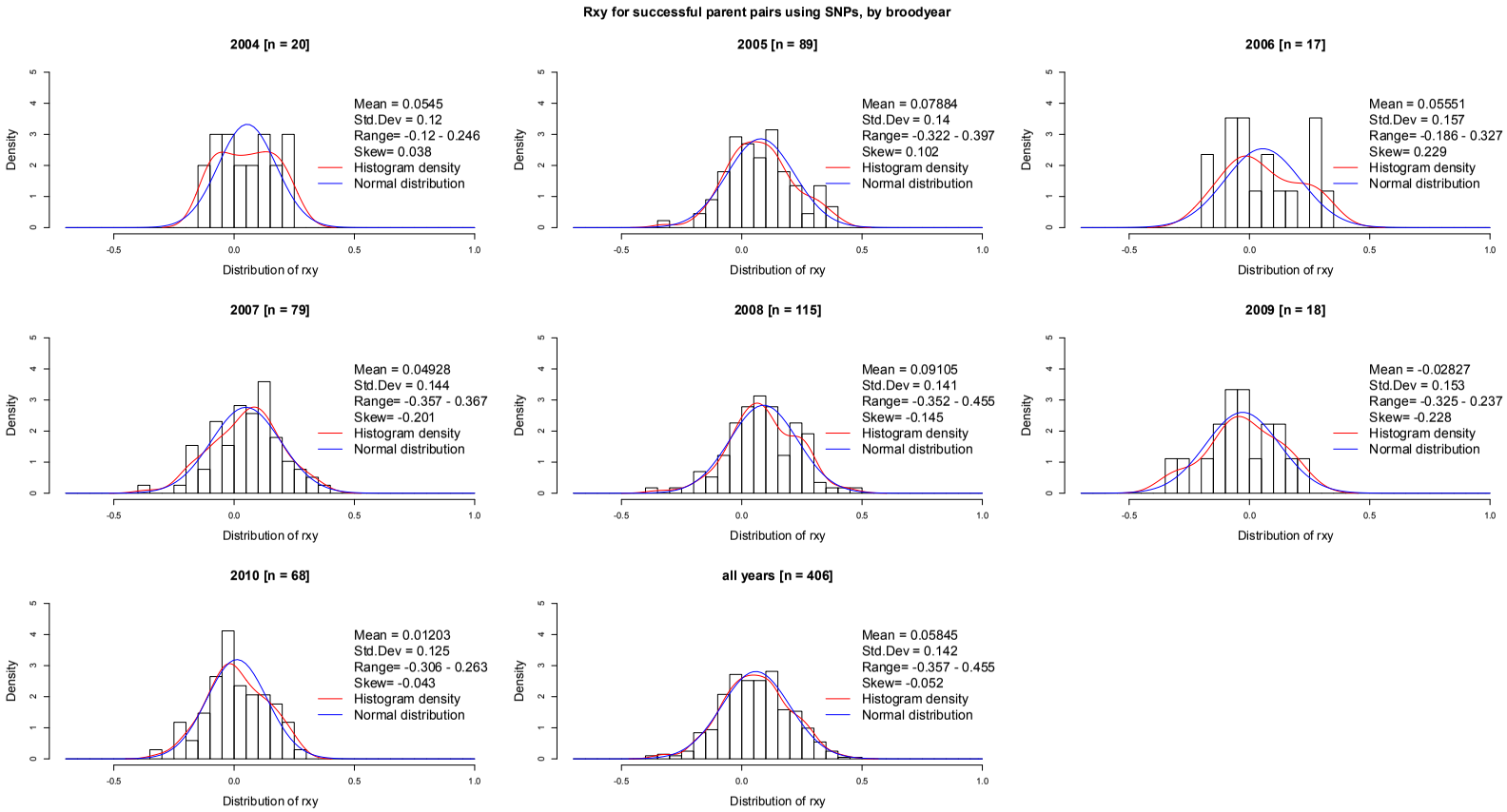

Figure 2.11: The distribution of relatedness generated with SNP genotypes for successful parent pairs (pairs with one or more adult offspring that returned to the hatchery). Mean, standard deviation (Std. Dev.), range and skew are reported.

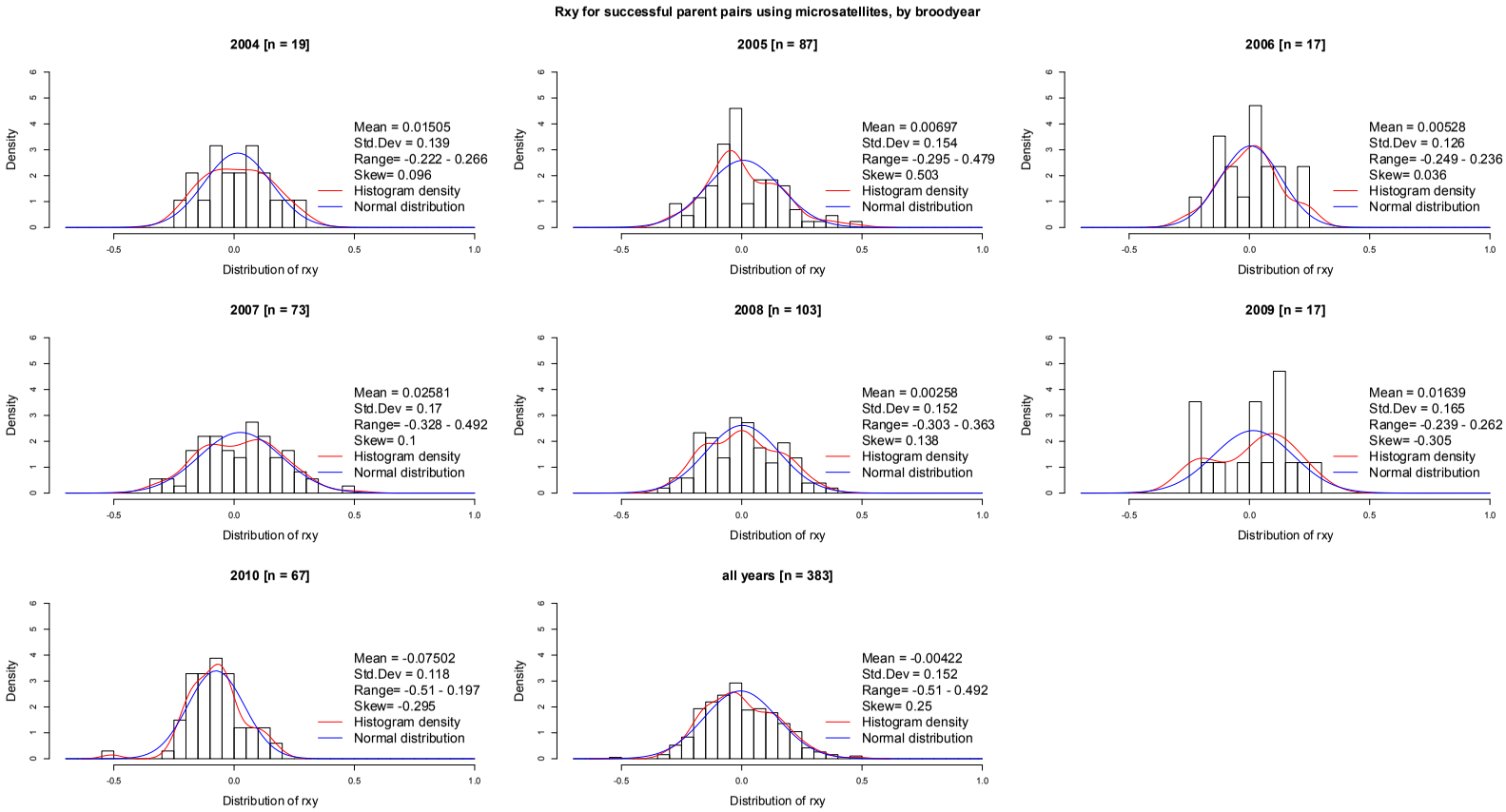

Figure 2.12: The distribution of relatedness generated with microsatellite genotypes for successful parent pairs (pairs with one or more adult offspring that returned to the hatchery). Mean, standard deviation (Std. Dev.), range and skew are reported.

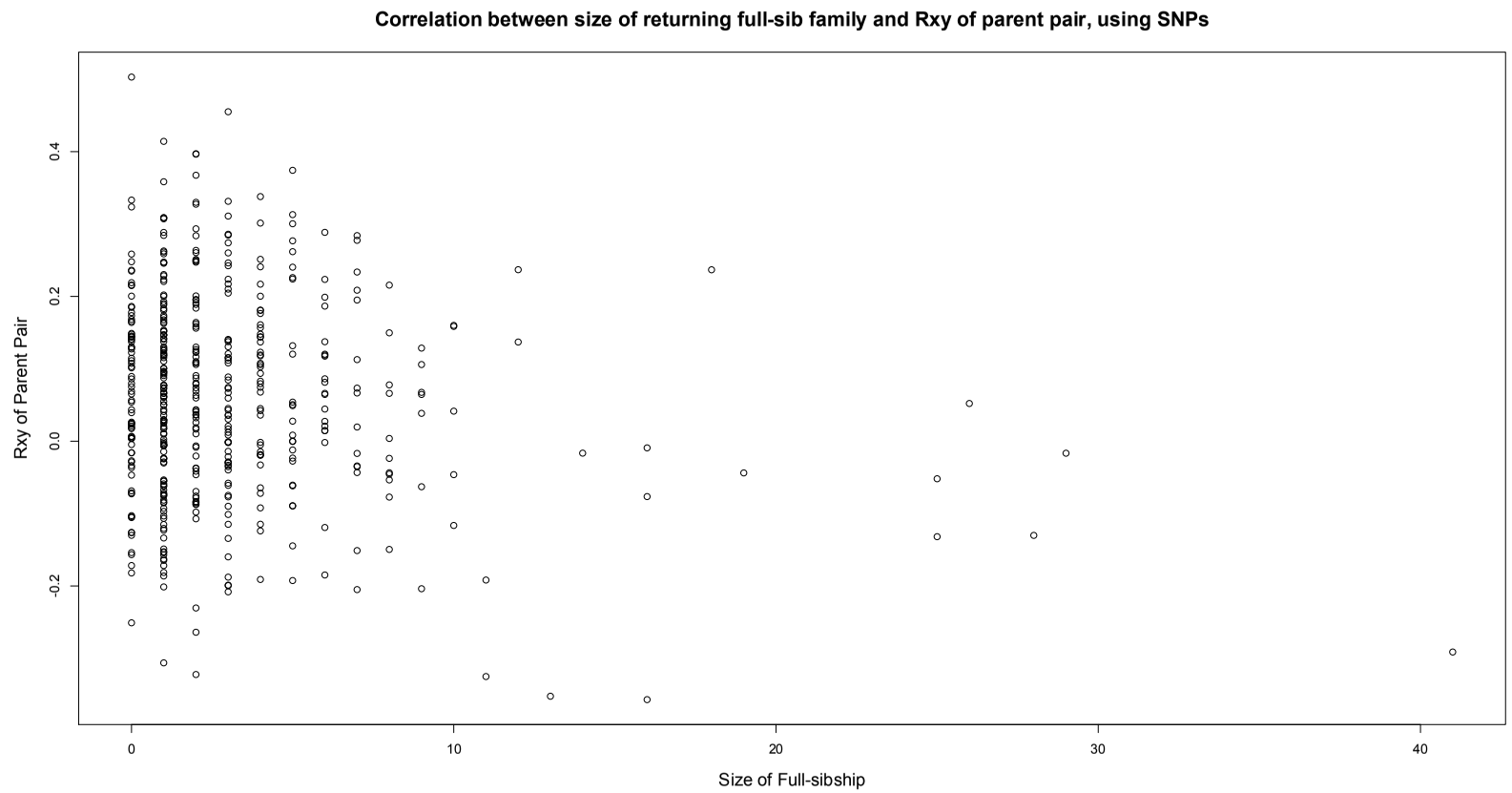

Figure 2.13: Correlation between the number of offspring that returned to the hatchery over all years, and the value of Rxy (generated by SNPs) between a parent pair.

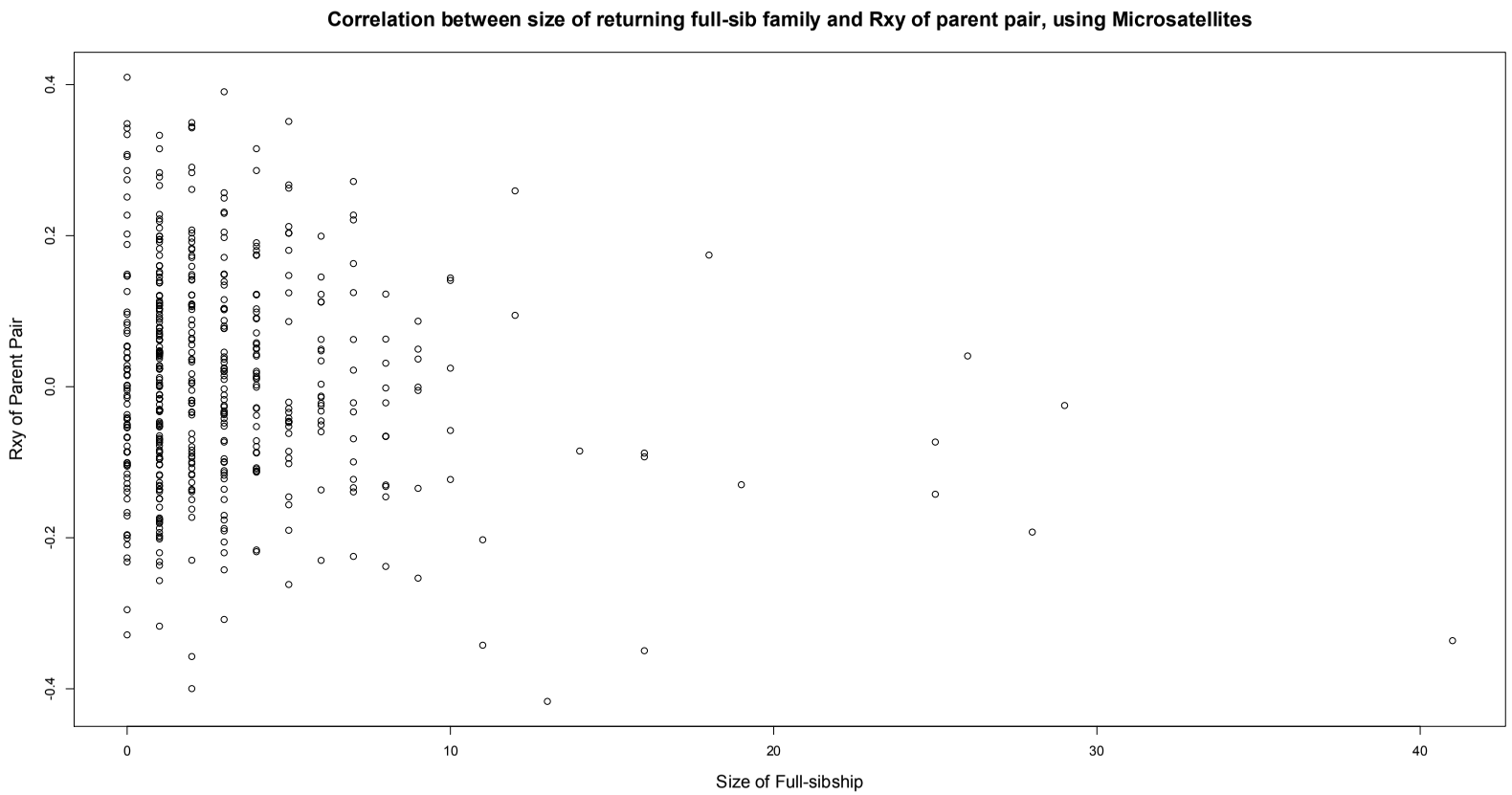

Figure 2.14: Correlation between the number of offspring that returned to the hatchery over all years, and the value of Rxy (generated by microsatellites) between a parent pair.

The parentage analysis detected a total of 1,381 full siblings and 333 half siblings (paternal: 188, maternal: 145) in the entire data set, allowing for the comparison of SNP- and microsatellite-generated Rxy values between the two known relationship types. Data for both marker types was approximately normally distributed and, for full siblings, although the mean value of Rxy for microsatellites (0.5138; range = -0.131 - 0.899) was greater than the mean for SNPs (0.5092; range = 0.075 - 0.868), these values were similar and close to what would be expected for full-siblings. For half siblings, however, the mean for the microsatellite-generated Rxy values (0.3351; range = -0.160 - 0.732), was greater than the mean from the SNP genotypes (0.1992; range = -0.206 - 0.639), and both differed from the expected value for half siblings. Since a large proportion of the half-sibs consisted of the super-jack's contribution (69 offspring with two females), this meant that the majority (1,148 of 1,447 or 79.3%) of all half-sib pairwise comparisons were generated from this one male. Because of this, distributions of Rxy for the two relationship types were examined again with this individual's contribution to both the full and half siblings removed. (This meant that 1,097 of 5,765, or 19% of all pairwise comparisons for the full siblings were removed.) Again, the mean values of Rxy generated by either marker type for the full-siblings were similar. However, this time the mean of SNP generated Rxy values (0.5166; range = 0.075 - 0.868) was larger than the mean of the microsatellite generated values (0.4934; range = -0.159 - 0.893). For the half siblings, removing the super-jack's contribution brought the mean values for both locus types closer to 0.25, the expected value for this relationship type, although in this case the mean value for SNPs (0.2629; range =

-0.142 - 0.639) was larger than that for microsatellites (0.2588; range = -0.167 - 0.612). For all distributions, microsatellite data yielded a broader range of Rxy values and a greater degree of standard deviation (range = 0.138 - 0.162) than did the SNP data (range = 0.119 - 0.137). Also, skew was negative in all distributions except for the SNP values from all half siblings (0.101). However, these values were more negative in the microsatellite distributions (range = -0.371 - -0.157) than in those of the SNPs (range = -0.284 - -0.105; Figure 2.15 and Figure 2.16).

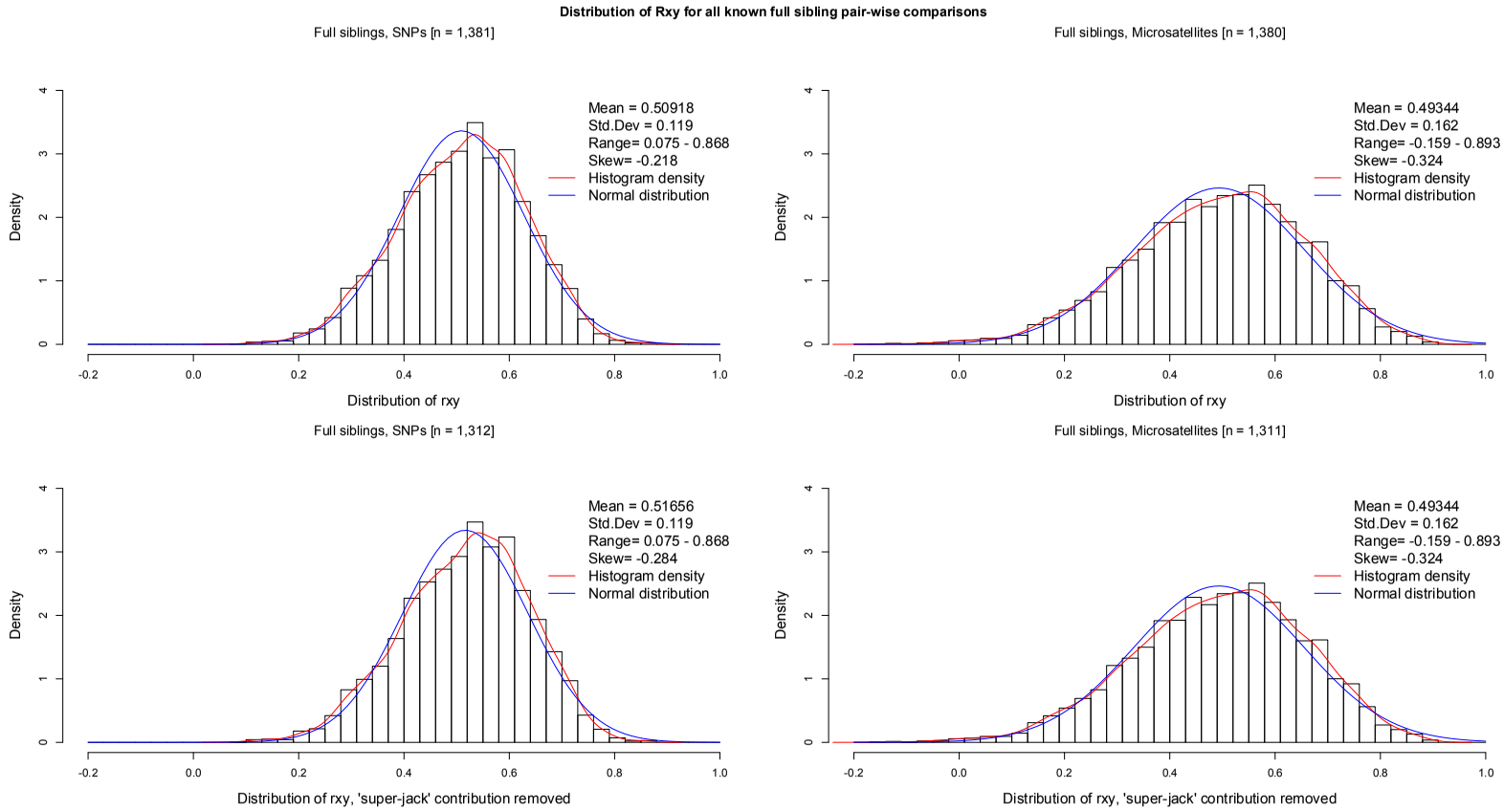

Figure 2.15: Distribution of relatedness generated with SNP and microsatellite genotypes for all full siblings (n) recovered in the parentage analysis. Distributions of full siblings are plotted both with and without the super-jack's contribution (69 offspring with two females) for both marker types. The mean, standard deviation (Std. Dev.), range and skew are reported.

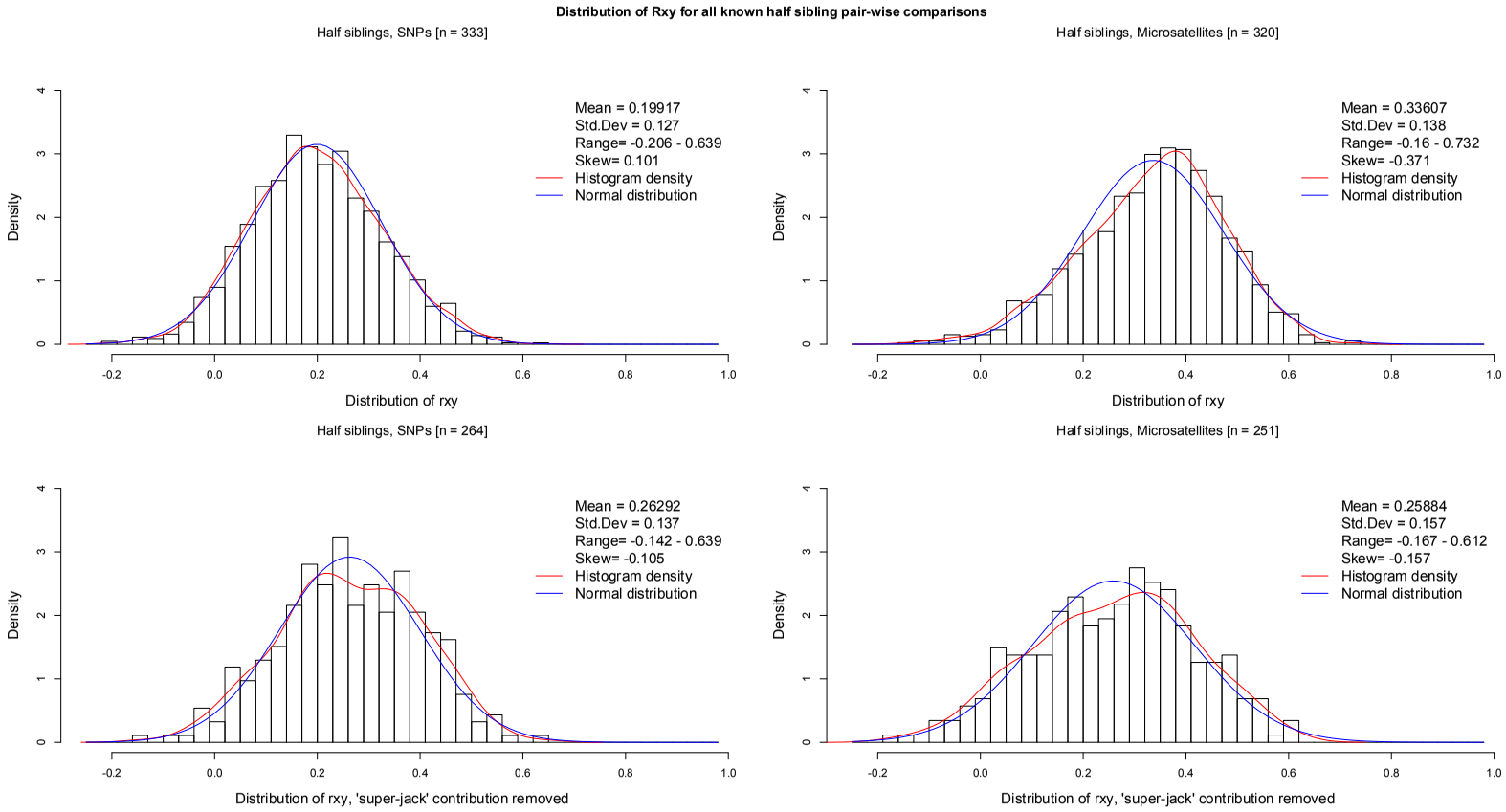

Figure 2.16: Distribution of relatedness generated with SNP and microsatellite genotypes for all half siblings (n) recovered in the parentage analysis. Distributions of half siblings are plotted both with and without the super-jack's contribution (69 offspring with two females) for both marker types. The mean, standard deviation (Std. Dev.), range and skew are reported.

The presence of a large group of individuals of known ancestry (full- and half-sibs) also allowed for the opportunity to quantify and compare the proportion of putative type II errors (related individuals being classified as unrelated) generated from either of the two marker types. For all full and half siblings, the proportion of pairs that achieved Rxy values less than 0.25 and less than 0.125 were calculated with both the SNP and microsatellite data, with and without the super-jack's contribution. An Rxy value of 0.125 is indicative of a first cousin relationship and is used as the cutoff value when determining potential mates in the hatchery.

Error frequencies generated within the two marker types for the combined group of full siblings and half siblings were plotted at given values of Rxy (-0.2 to 0.5), both with and without the super-jack's contribution to the half siblings (Figure 2.17, Figure 2.18 and Figure 2.19). Points on the plots indicate the proportion of Rxy values that fall below given Rxy cutoff points. The plotted error frequencies visually illustrate the differences between the two sets of markers. When all siblings were considered, SNP-generated Rxy estimates had greater frequencies of error than those generated from microsatellites. However, when the super-jack's half-sib offspring were removed from the analysis, SNP-generated Rxy estimates resulted in the lowest error frequencies between the two marker types.

The SNP-generated values of Rxy had smaller error rates than for microsatellites when estimating relatedness for the known full siblings, as 1.68% (97 of 5,765 pairs) of values fell below 0.25 and 0.09% (5 of 5,765 pairs) of values fell below 0.125, and with the microsatellites 5.43% (287 of 5,282 pairs) fell below 0.25 and 1.04% (55 of 5,282

pairs) fell below 0.125. However, when estimating values for the half siblings, the SNP genotypes generated larger error rates than did the microsatellite genotypes. For the SNP Rxy values, 66.3% (960 of 1,447 pairs) fell below 0.25 and 28.6% (414 of 1,447 pairs) fell below 0.125, while for the microsatellites, 25.6% (337 of 1,315 pairs) of the values fell below 0.25 and 7.83% (103 of 1,315 pairs) fell below 0.125. Removing the super-jack's contribution had little effect of the proportion of error in the full siblings for both the SNP- and microsatellite-generated data. Proportions were still smaller for the SNP-generated values of Rxy where 1.61% (75 of 4,668 pairs) fell below 0.25 and 0.11% (5 of 4,668 pairs) below 0.125, than for the microsatellite values where 7.48% (346 of 4,626 pairs) fell below 0.25 and 1.56% (72 of 4,626 pairs) fell below 0.125. For the half-sibs however, the proportion of SNP-generated errors decreased while the proportion of microsatellite-generated errors increased. With the SNPs, 47.9% (148 of 309) of Rxy values were less than 0.25 and 15.9% (49 of 309 pairs) were less than 0.125, while with the microsatellites 45.7% (133 of 291 pairs) were less than 0.25 and 20.3% (59 of 291 pairs) were less than 0.125. For full- and half-sibs combined, 5.81% (419 of 7,212 pairs) of all SNP generated values of Rxy fell below 0.125, while 2.40% (158 of 6,597 pairs) of all microsatellite-generated values fell below 0.125. When the super-jack's contribution was removed, 1.08% (54 of 4,977 pairs) of the Rxy values were less than 0.125 with SNP genotypes and 2.66% (113 of 4,917 pairs) of the values were fell below 0.125 with microsatellite genotypes (Table 2.7).

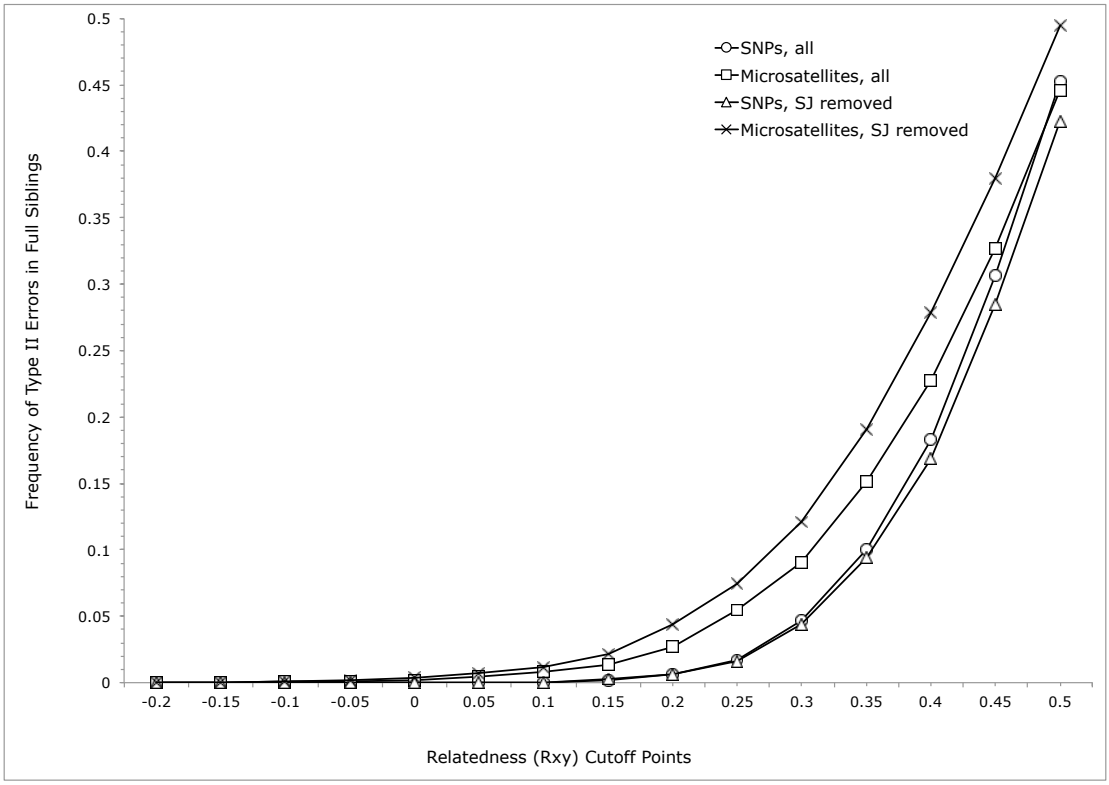

Figure 2.17: Frequency of type II errors in the full siblings (related individuals classified as unrelated) for SNP- and microsatellite-generated values of relatedness (Rxy), both with and without the super-jack's contribution.

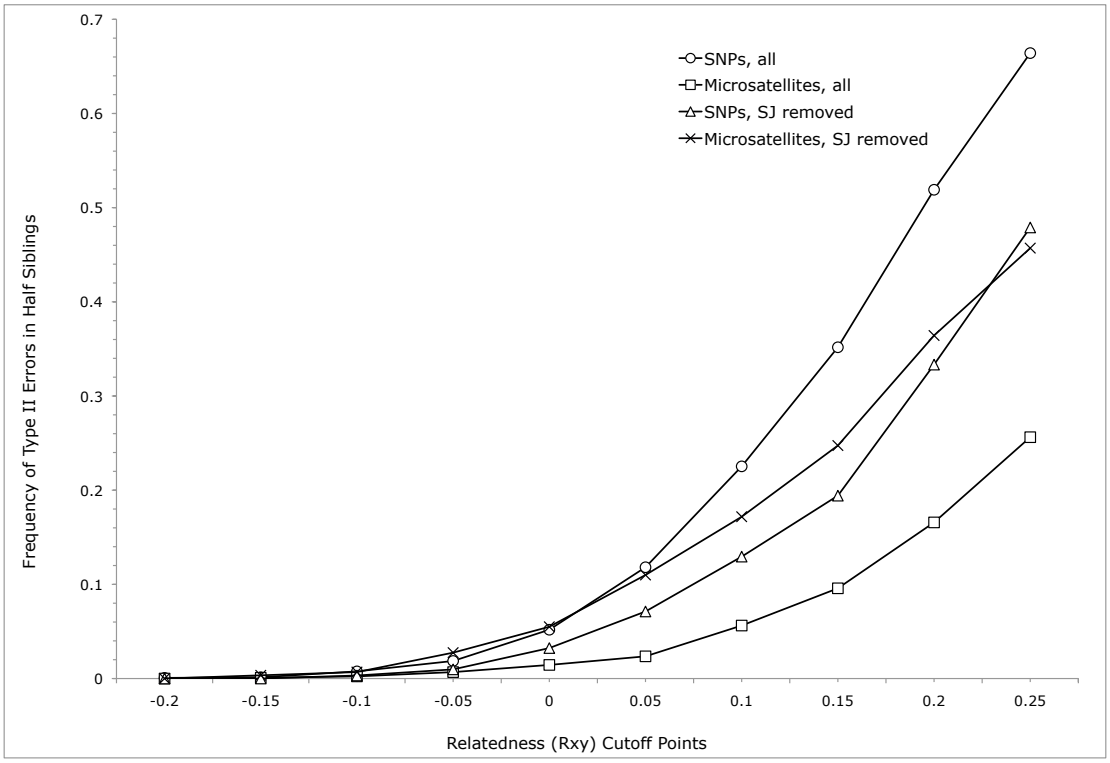

Figure 2.18: Frequency of type II errors in the half siblings (related individuals classified as unrelated) for SNP- and microsatellite-generated values of relatedness (Rxy), both with and without the super-jack's contribution.

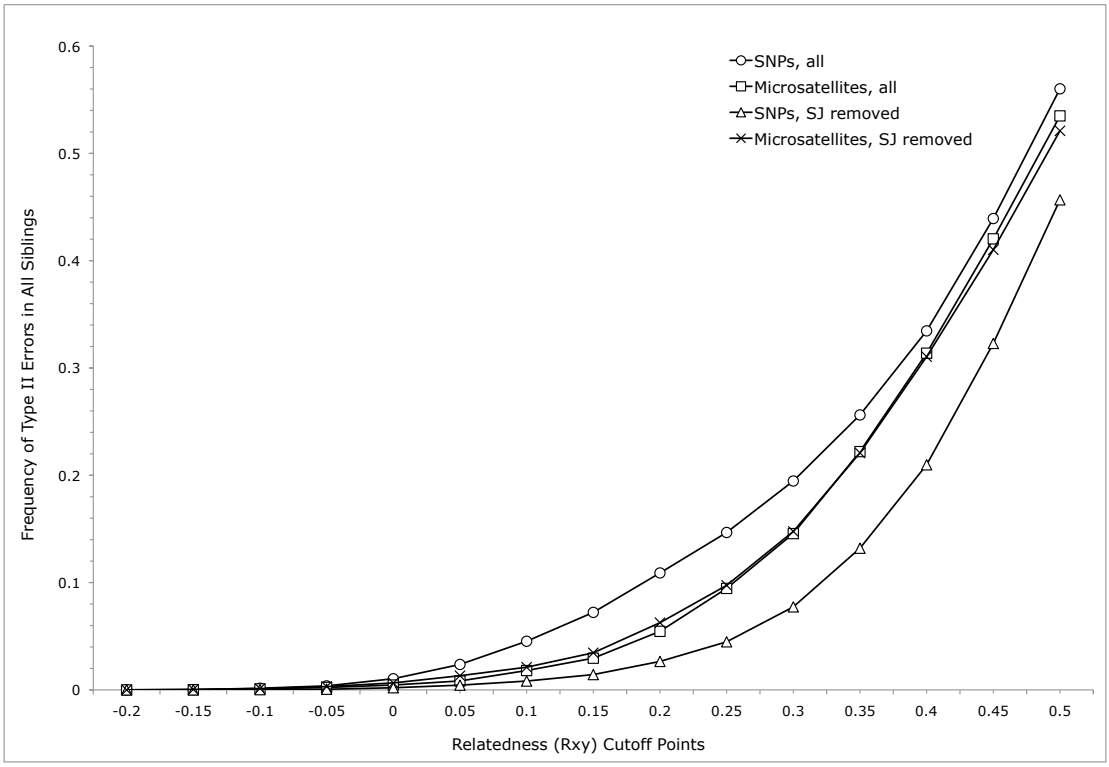

Figure 2.19: Frequency of type II errors in all siblings (related individuals classified as unrelated) for SNP- and microsatellite-generated values of relatedness (Rxy), both with and without the super-jack's contribution.

Table 2.7: Relatedness (Rxy) type II error rates generated with SNP and microsatellite markers, calculated with 0.25 and 0.125 as cutoff points (Figure 2.17, Figure 2.18 and Figure 2.19). Comparisons are made for full siblings, half siblings and the pooled group of siblings between the two marker types both with and without the super-jack's contribution.

	All pairs		**'Super-jack' removed**	
Misidentification type	**93 SNP loci**	**16 Msat loci**	**93 SNP loci**	**16 Msat loci**
Proportion misidentified using 0.25 relatedness as cutoff point				
Full-sibs ranked as unrelated	0.0168	0.0543	0.0161	0.0748
Half-sibs ranked as unrelated	0.6634	0.2563	0.4790	0.4570
All siblings ranked as unrelated	0.1466	0.0946	0.0448	0.0974
Proportion misidentified using 0.125 relatedness as cutoff point				
Full-sibs ranked as unrelated	0.0009	0.0104	0.0011	0.0156
Half-sibs ranked as unrelated	0.2861	0.0783	0.1586	0.2027
All siblings ranked as unrelated	0.0581	0.0240	0.0108	0.0266

Since family size ranged from 2 to 41 offspring for full siblings, and 2 to 69 offspring for half siblings, SNP genotypes for both sibling groups were simulated in order to examine the distributions of Rxy from sibling pairs with no more than 2 offspring per family. Values of Rxy were calculated in ten groups of simulated full and half siblings, each consisting of 500 individuals. For the full siblings, Rxy was approximately normally distributed. The mean Rxy for all individuals combined was 0.4936 (range = 0.034 - 0.862), while across groups the mean ranged from 0.4859 to 0.5017. Skew was negative in all groups except one (skew = 0.122) and ranged from -0.612 to -0.078, while over all simulated pairs it was -0.297. For the half siblings, values were approximately normally distributed and the mean Rxy for all individuals combined was 0.2490 (range = -0.105 - 0.608). Across groups, the mean ranged from 0.2394 to 0.2600. Skew was negative in all but one group (skew = 0.025), and ranged from -0.282 to -0.017, while over all individuals combined skew was -0.141 (Figure 2.20).

A linear correlation between microsatellite- and SNP-generated Rxy values was examined for each brood year and plotted for all collection years combined (Figure 2.21). Significant positive correlations ($P < 0.001$) were observed in all years and over all years combined, but variability was high. Values of R^2 ranged from 0.004 in 2007 ($F_{1,77419}$ = 319) to 0.093 in 2006 ($F_{1,10876}$ = 1,110) and equaled 0.030 ($F_{1,530540}$ = 16,540) over all years, while the slope of the correlation lines (b) ranged from 0.062 in 2007 to 0.299 in 2006 and was 0.167 over all years combined (Table 2.8). Because the slopes for these correlations are likely driven by the presence of a large proportion of siblings, the same comparisons were performed, but for all pedigree confirmed siblings both with and

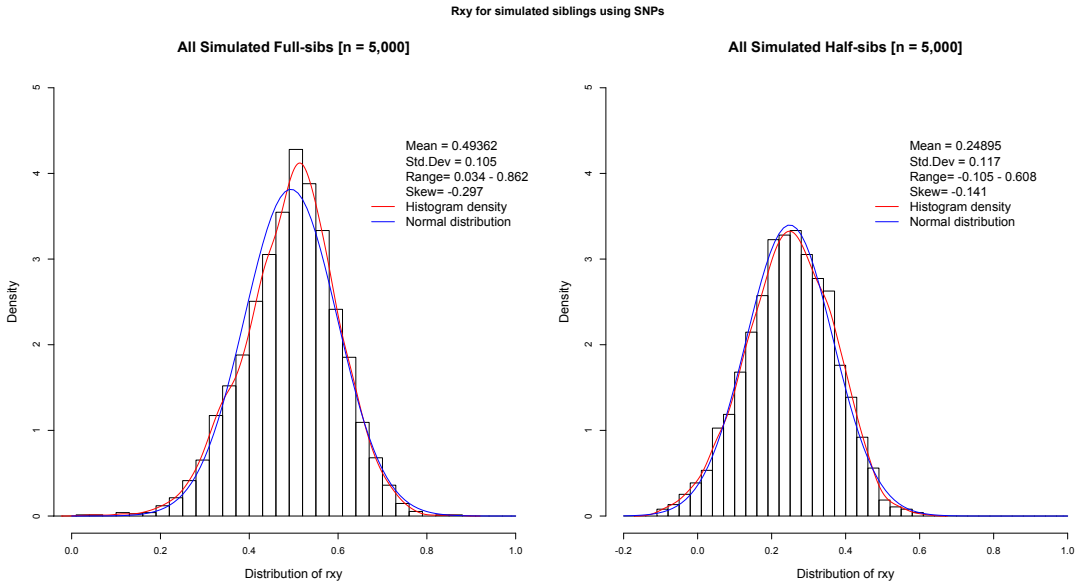

Figure 2.20: Distributions of relatedness (Rxy) for 10 groups of simulated full and half siblings combined. Each group consists of 500 individuals with two full siblings and two half siblings per parent pair, making a total of 5,000 individuals combined per sibling group. As the data were approximately normally distributed, the mean, Std. Dev. (standard deviation), range and skew are reported.

without the super-jack's contribution. In this case, a significant negative correlation was found for all comparisons except those that included both full- and half-sibs combined. Of all comparisons, b was the largest (0.270) with the strongest correlation ($F_{1,5992}$ = 427.1, R^2 = 0.066, P < 0.001) for all siblings combined (Figure 2.22). When the super-jack's contribution was removed, variability increased ($F_{1,4599}$ = 46.25, R2 = 0.010, P < 0.001) and the slope decreased to 0.078. As before, the presence of the super-jack's offspring had less effect on the full siblings than on the half siblings when compared separately. The slope was negative but close to zero when all full siblings were compared (-0.027) and variability was high ($F_{1,4687}$ = 6.015, R^2 = 0.001, P < 0.05) as well as when the super-jack's contribution was removed (b = -0.031, F1,4344 = 7.612, R^2 = 0.002, P < 0.01). When all half siblings were compared, the slope was negative (-0.069) and variability was high ($F_{1,1303}$ = 7.387, R^2 = 0.005, P < 0.01). When the super-jack's contribution was removed, the slope became more negative (-0.150) and variability decreased ($F_{1,253}$ = 7.588, R^2 = 0.025, P < 0.01; Table 2.9).

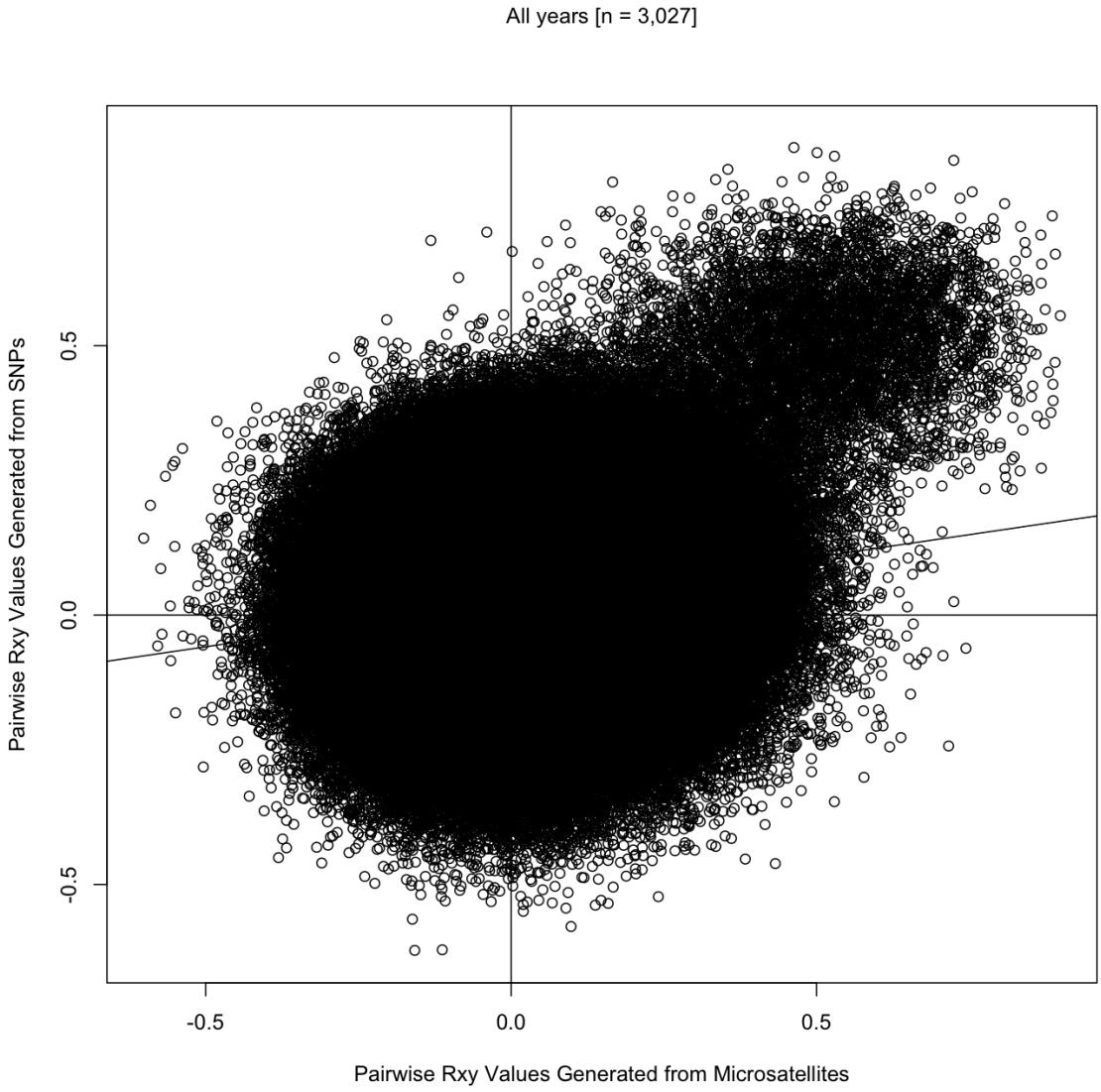

Figure 2.21: Linear correlation of all possible pairwise values of Rxy generated from microsatellites on the same pairs of Rxy generated from SNPs over all years (2004 - 2012) combined.

Table 2.8: The slope of the correlation line (b) between microsatellite- and SNP-generated values of Rxy for each brood year (2004 - 2012) and over all brood years combined (Figure 2.21). The goodness of fit (R^2) and standard error (SE) are also reported. ***Significant at $p < 0.001$.

Brood year	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**All**
b	0.082***	0.133***	0.299***	0.062***	0.208***	0.164***	0.121***	0.232***	0.193***	0.167***
R^2	0.007	0.019	0.093	0.004	0.050	0.030	0.014	0.056	0.047	0.030
SE	0.141	0.138	0.147	0.138	0.137	0.140	0.144	0.140	0.144	0.143

Table 2.9: The slope of the correlation line (b) between microsatellite versus SNP generated values of Rxy in all siblings combined, full siblings and half siblings, with and without (rmv) the super-jack's contribution (Figure 2.22). The goodness of fit (R^2) and standard error (SE) are also reported. ***Significant at $p < 0.001$, **Significant at $p < 0.01$, *Significant at $p < 0.05$.

Siblings	**All**	**All (rmv)**	**Full**	**Full (rmv)**	**Half**	**Half (rmv)**
b	0.270***	0.078***	-0.027*	-0.031**	-0.069**	-0.150**
R^2	0.066	0.010	0.001	0.002	0.005	0.025
SE	0.171	0.133	0.120	0.120	0.126	0.132

Delta Rxy, or the difference in values of Rxy between microsatellites and SNPs, was calculated by subtracting SNP-generated values of Rxy from the corresponding microsatellite values. The distributions of delta Rxy were plotted for the two sibling groups, both with and without the super-jack's contribution. For all four distributions, the data were approximately normally distributed. For all full siblings, mean delta Rxy was -0.0118 (range = -0.775 - 0.632) and for full siblings with the super-jack's contribution removed, the mean was -0.0248 (range = -0.769 - 0.632). For all half siblings, the mean was 0.1400 (range = -0.517 - 0.814), and when the super-jack's contribution was removed, the mean was -0.0125 (range = -0.539 - 0.558). Skew was small and negative in all comparisons (range = -0.319 - -0.103) except that of the half siblings with the super-jack's contribution removed, where it was small but positive (0.052; Figure 2.23).

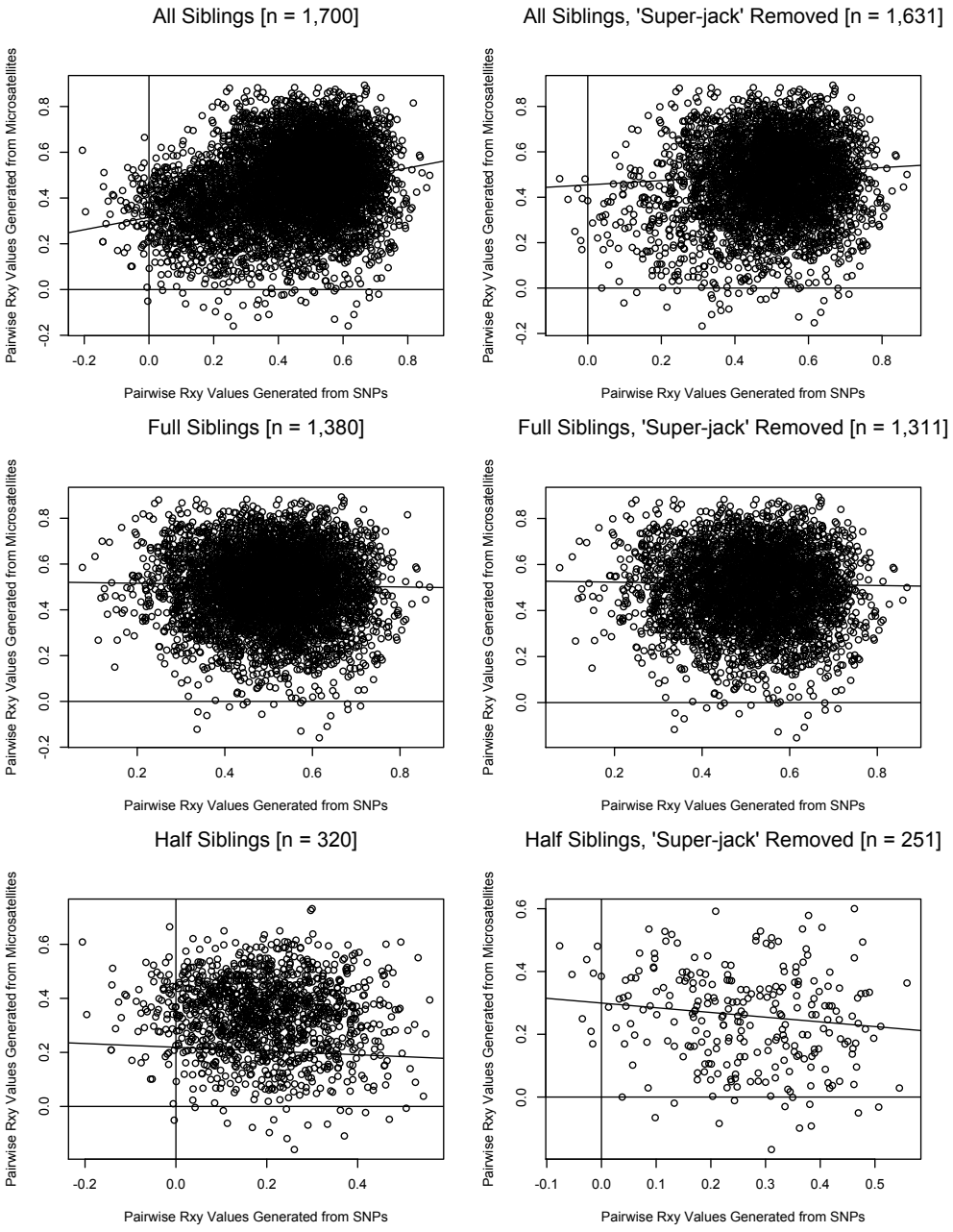

Figure 2.22: Linear correlation of pairwise values of Rxy in siblings generated from microsatellites and SNPs. Independent comparisons were made for all siblings combined, all full siblings and all half siblings both with and without the super-jack's contribution.

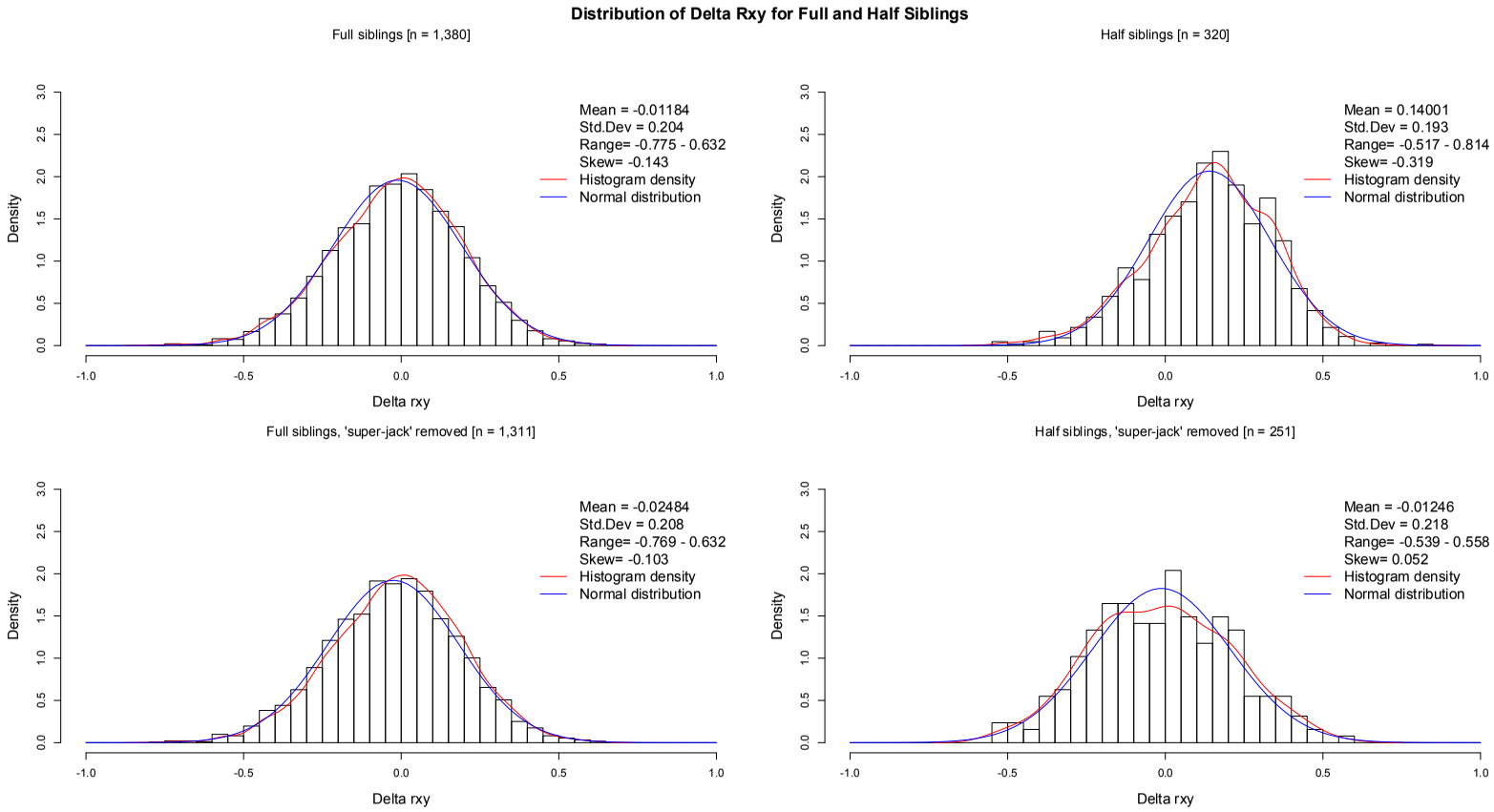

Figure 2.23: Distribution of delta Rxy, or the difference in values of Rxy generated by microsatellites and SNPs. As values are approximately normally distributed, the mean, standard deviation (Std. Dev.), range and skew are reported.

2.4 Discussion

This study confirms the power and utility of pedigree-based intergenerational genetic tagging using a multi-locus panel of SNPs for parentage inference in a medium sized hatchery population of coho salmon, a species facing extinction in much of its native range. By genotyping nearly all broodstock over a period of 9 years, with a panel of 96 SNP markers and utilizing novel and efficient algorithms for parentage reconstruction, I was able to recover offspring as they returned to the hatchery to spawn at two and three years of age. Offspring were assigned to parents with high accuracy, which could be confirmed for most of the assignments by the hatchery recorded mate pairs. I was then able to use the reconstructed pedigrees to investigate variance in reproductive success, age structure of cohorts as well as adult offspring returning to spawn, and the potential for heritability of length at maturity. Additionally I was able to perform an in depth comparison of relatedness values, generated from SNPs and microsatellites, for the entire sample set as well as for pedigree-confirmed siblings.

2.4.1 Pedigree Reconstruction

The resulting pedigrees were reconstructed with high confidence, as denoted by the low FDR scores and high posterior probabilities. The accuracy of these pedigrees was also confirmed by hatchery documentation of mated pairs. The proportion of offspring that assigned to parent pairs was approximately 83% when only accounting for offspring years in which complete sampling of the parental generations took place (2008 - 2012).

The majority of unassigned offspring is most likely explained by those parents that were excluded from the analysis due to excessive missing data or because samples for certain individuals were not received. The scaled assignment estimates for offspring from 2008 to 2012, indicate that the expected proportion of assigned offspring would have increased to approximately 92% had parents (from 2005 to 2010) not been excluded from the data set due to excessive missing data or lack of sampling. Additionally, these proportions were calculated by dividing the total number of assignments by the total number of LM clipped and 'unknown' fish (individuals lacking hatchery mark information) combined. It is possible that some fraction of these 'unknown' fish are not actually fish produced at Iron Gate Hatchery (LM clipped) but instead are of natural origin or are strays from the Trinity or Rogue River. The inclusion of non Iron Gate fish in the total returns would drive this calculation down because the parents of these fish would most likely not have been spawned in Iron Gate hatchery. It is also possible that a small portion of fish spawned outside of Iron Gate and their offspring were collected at the hatchery, which would also skew calculations as the parents would not be in the database.

2.4.2 Age Structure of Returning Adults

Only two- and three-year olds were recovered in this analysis. Four-year old adult coho salmon are a common occurrence in the more northern reaches of the species' range, and in fact make up the predominant age class in southeast and central Alaska, where they spend two years in freshwater (as opposed to one) before smolting and migrating to sea (Weitkamp *et al.* 1995). Although this extended life history is generally

not associated with populations in California (Shapovalov and Taft 1954; Weitkamp *et al.* 1995), Bell and Duffy (2007) observed the presence of two-year old out-migrants in Prairie Creek, a coastal stream in northwest California situated within both Redwood State and National parks. Although their study was conducted on juveniles migrating to the ocean (not on adults returning to spawn), it is possible that a portion (if not all) of these two-year old smolts returned to their natal stream as four-year olds based on this older age at out-migration. Although the mouth of the Klamath River is located approximately 16 km north of that of Prairie Creek, and both river systems are located within the SONCC ESU, four-year old adults returning to IGH were not recovered in this analysis.

In all three cohorts (and throughout the total analysis) only males returned as precocious two-year olds, which is consistent with coho salmon life history traits in the more southern reaches of the species' native range (Shapovalov and Taft 1954; Weitkamp *et al.* 1995; Lestelle 2007). The proportion of age two and age three adults that returned to the hatchery differed significantly between the 2007, 2008 and 2009 cohorts. The proportion of jacks was the smallest in the 2007 cohort making up almost 2% of all individuals, while proportions increased to approximately 33% in the 2009 cohort. Although three-year olds were the predominant age class in all three cohorts, in the 2009 cohort, jacks outnumbered age-three males by approximately 61%. Additionally, while age-three fish were the dominant age class in the 2010 and 2011 return years, jacks made up approximately 64% of all individuals, and nearly 85% of all males, in the 2012 return year. It is important to note that the 2009 and 2010 spawn years mark

the first instances in which jacks were purposely spawned at the hatchery. In 2009, 5 of the 16 fathers that recovered offspring were jacks, while in 2010, 9 of the 60 fathers were jacks. As precocity has been found to have a heritable component in coho and Chinook salmon (Iwamoto *et al.* 1984; Hankin *et al.* 1993; Appleby *et al.* 2003), it is possible that the large proportion of jacks in the 2009 cohort and the 2012 return year is a reflection of those that were mated in the 2009 and 2010 spawn years.

However, despite the underlying heritable component believed to influence age at maturity, in the southern portion of the species' range the propensity for male coho salmon to mature at age two, a characteristic that tends to decrease with increasing latitude (Drucker 1972), is an outcome thought to be driven primarily by the quality and productivity of the freshwater habitat (Hager and Noble 1976; Young 1999; Lestelle 2007). Young (1999) found that increased freshwater growth rates and large smolt size, resulted in a greater occurrence of male coho returning at age two. Coho salmon are reared at Iron Gate Hatchery until they are released as smolts. Since several studies have shown that hatchery reared fish tend to grow faster and larger in a shorter period of time than those reared in the wild, presumably owing to a consistent food supply (pellets as opposed to naturally sourced food) and relaxed selective pressures (i.e. relaxed competition for food as well as a decreased need for predator avoidance; Chittenden *et al.* 2010; Christie *et al.* 2011), the large proportions of jacks in the data set is perhaps not such a surprising result. In fact, Iron Gate Hatchery records dating back to 1993 report large proportions of jacks (based on visual identification) in prior cohorts, the largest proportion occurring in the 1996 cohort, where jacks made up approximately 51% of

all individuals and approximately 64% of all males (HSRG 2012). As males of lengths 580 mm or smaller have been excluded from spawning in the past at the hatchery, the likelihood that precocious individuals were spawned accidentally is minimal based on the clearly separated size distributions between age-two males and age-three males and females shown in this analysis (Figure 2.6). However, when considering the numbers of jacks returning per year to the hatchery, 2012 marks the year in which the greatest proportion of precocious males was recorded since recording began in 1993. From 1993 to 2009, the highest proportions of jacks that returned occurred in 2000, where they made up 47% of all individuals and 68% of all males (HGMP 2012). In this case however, three-year-old adult fish were still the dominant age class unlike in 2012 when precocious males dominated.

The small proportion of three-year olds that returned to the hatchery in 2012 is likely a reflection of the poor return year in 2009, where only 40 individuals were spawned. Even though these spawners produced large families, the total number of age three offspring was minimal in comparison to prior years, likely due to limited egg production (lowest since 1993; HSRG 2012), as only 20 females were available to spawn. As both environmental and genetic factors play a role in age-at-maturity for coho salmon (Appleby *et al.* 2003), the extent to which each factor is contributing is not totally clear. Ultimately, through continued monitoring of this system using pedigree analysis, separating the genetic component of jacking from the environmental component may soon be possible.

2.4.3 Family Size and Relative Reproductive Success

Although from 2004 to 2009, management guidelines at Iron Gate hatchery specified one-to-one matings between males and females, based on the numerous half siblings recovered in the analysis, these guidelines were not always adhered to. In the 2005 through 2009 brood years, the parentage analysis recovered 14 males (out of 301) that were spawned with two females (in one case three females) and 13 females (out of 303) that were spawned with two males (two multiply spawned males, and two multiply spawned females were recovered in 2004, however not all broodstock were collected for genetic analysis so the total number of multiply spawned individuals from this year cannot be estimated). This means that 4.7% of males were spawned at a male-to-female ratio of one-to-two and that 4.3% of females were spawned at a female-to-male ratio of one-to-two. Despite this small diversion from the desired mating scheme in these years, reproductive success was near equal (due to the similar promiscuous mating scheme used for both sexes), with males having a slightly larger coefficient of variance most likely resulting from the super-jack's 69 offspring with two females.

In 2010, hatchery guidelines were modified to include two-to-one matings between males and females when female returns were low (HGMP V.10 2013). This means that female spawners are to be crossed with one or two males, while males are to be crossed with only one female. Although 9 of the 59 female spawners (15.3%) were spawned with more than one male, 7 of the 60 (11.7%) male spawners recovered in the analysis had more than one female mate, meaning that hatchery practices veered away

from the desired mating scheme in this year to a small extent. Despite this, the relative reproductive success between sexes was still near equal.

The purpose of a promiscuous mating scheme, where multiple males mate with one female, preferably only once, is to increase family numbers as well as genetic diversity, and to decrease the chance of inbreeding relative to single pair or factorial mating (Pearse and Anderson 2009). As the results of this analysis indicate that overall family size is large in this population, it is necessary to increase family number in order to reduce family size. A population consisting of a few large families increases the likelihood of inbreeding among individuals, as large families have an increased chance of interbreeding (Christie *et al.* 2011). Because hatchery fish tend to have negative impacts on nearby natural populations of the same species, it is important to focus on improving the genetic integrity of hatchery populations or stocks. In the Klamath, IGH fish have been known to stray and spawn in the Shasta River, a nearby population unit in the Klamath basin (HGMP V.10 2013). This natural population (as well as all others in the Klamath basin) is also ESA listed, and is experiencing greatly reduced numbers. IGH fish threaten the genetic integrity of the Shasta population through the risk of introgressing their more genetically depauperate genes into these natural stocks. For this reason it is especially important to improve hatchery practices in order to produce more genetically fit fish, so that they will have a lesser negative impact on their wild counterparts. The results of this analysis show that monitoring through pedigrees will allow for routine evaluation of spawned individuals, distributions of family size, and the potential for inbreeding at the hatchery, in an accurate and informative way.

2.4.4 Heritability of Length-at-Maturity

Morphological traits such as body size and length-at-maturity have been found to have strong heritable components in coho salmon, as well as in other salmonids (Ricker 1981; Funk *et al.* 2005), however, the length-at-maturity results in this analysis were somewhat inconclusive. Significant relationships occurred when regressions of midparent length on all male offspring, mother length on all offspring (both sexes), and mother length on all male offspring were made. These results seem to indicate a maternal effect on offspring size, particularly for male offspring size. However, variability was also great in all correlations, based on the small R^2 values, suggesting that the environmental component influencing an individual's size had a greater influence than parental size. This is perhaps not entirely surprising as this is a hatchery population where it is assumed that selective pressures have been relaxed. It has been shown that larger juvenile coho salmon experience higher rates of freshwater survival as well as increased marine survival (Bilton *et al.* 1982; Quinn and Peterson 1996). However in the hatchery, owing to consistent food supply and domestication selection, it is likely that most juveniles achieve large size before migrating to sea, so heritability estimates could be confounded by the hatchery environment as well as small sample size.

2.4.5 Explorations in Relatedness

Inbreeding has been shown to have deleterious effects on salmon fitness and survival, especially in the context of hatcheries and captive breeding programs (Wang *et al.* 2002; Conrad *et al.* 2013). In the wild, salmon presumably detect or smell chemical

cues that identify other siblings or other unrelated individuals (Quinn and Busack 1985; Olsen 1999) in order to avoid matings between close kin. There is also evidence that females select mates non-randomly based on mate compatibility at the major histocompatibility complex (Bernatchez and Landry 2003; Neff *et al.* 2008; Evans *et al.* 2012). However, in a hatchery setting the elements of sexual selection and kin avoidance are removed because mates are selected by hatchery staff, although certain preferred traits such as growth rate are often simultaneously selected for (Becker *et al.* 2013). In such situations, the broodstock is generally of unknown pedigree, which increases the risk of spawning related individuals, especially when run years have been severely reduced (Wang *et al.* 2002). Genetic parameters that estimate relatedness between putative breeding pairs have become an important means for reducing the occurrence of inbreeding in many hatchery populations (*Warm Springs Hatchery* , Conrad *et al.* 2013; *Scott Creek Hatchery* , Sturm *et al.* 2009; *Iron Gate Hatchery* , HSRG 2012).

The most striking result from the relatedness analysis is the effect that the super-jack's offspring had on the distributions of Rxy values in the half sibling group for both marker types. When the super-jack's offspring were included, SNPs underestimated values of Rxy (mean, 0.199), while microsatellites overestimated values of Rxy (mean, 0.336). Of the total 333 half siblings recovered through the pedigree analysis, 69 (20.7%) were derived from this one male, meaning that his offspring accounted for 79% (1,148 of 1,447) of all pairwise values generated with the SNP markers (78% or 1,024 of 1,315 for the microsatellite markers). When the super-jack's offspring were removed from the half sibling analysis, the mean Rxy values for both marker types were near

equal to 0.25 (0.263 with SNPs and 0.259 with microsatellites), the expected relatedness value for half siblings.

It is not completely clear why the inclusion of such a large group of half siblings shifts the distributed Rxy means in the negative direction for the SNP values and in the positive direction for the microsatellite values. It is however apparent that the overrepresentation of one individual's alleles greatly affects the way in which Rxy is calculated when using either marker type. The obvious difference between the two marker types is the number of alleles present at each locus; at the 16 microsatellite loci used in this analysis, 366 alleles have been identified in California coho (Conrad *et al.* 2013), while only 186 alleles have been identified at the 93 SNP loci, as these SNPs are biallelic. It could be that when a dataset is swamped by one individual's alleles, the greater allelic richness in the microsatellites infers a greater than expected mean pairwise Rxy value for the half-sibship sired from this one individual. Conversely, it could also be that the reduced allelic richness in the SNPs infers a smaller than normal mean Rxy value within the same group of individuals, because there aren't enough different alleles to more precisely discriminate the half sibling relationship.

Another explanation could be that the super-jack had an elevated internal heterozygosity in his SNP genotypes resulting from contaminated DNA, which then made it appear that he sired more offspring than is really the case. In our lab, we use an internal heterozygosity cutoff of 0.56 or more in order to detect DNA contamination when genotyping with SNPs. The super-jack's internal heterozygosity was 0.333, however, which is within the range of values expected for SNPs (between 0.16 and 0.56).

An alternative explanation could also be that the two females that spawned with the super-jack represent two of the least related individuals in the dataset, however their pairwise Rxy value was 0.0924, which is not representative of the smallest Rxy values in the dataset. It is however interesting to note that the Rxy values between the super-jack and each spawned female were very small (-0.1301 for female 1 and -0.2912 for female 2). Small Rxy values would generally be expected between a jack and a three-year old in this situation because they represent different cohorts and would likely not share a large degree of relatedness, however when examining relatedness between the super-jack and female 1 (female 2 was discarded from the microsatellite analysis due to excessive missing data), their Rxy value was larger (0.0422). Whether or not this difference between parental values of Rxy between the two marker types is large enough to affect the output of pairwise Rxy values in the half-sibs to the degree exemplified in this analysis is unclear, but it could be a possible explanation.

In terms of inbreeding avoidance, overestimating the degree of relatedness among potential spawners is preferable to underestimating relatedness, so as not to mistakenly categorize individuals as unrelated and ultimately spawn them when in fact they are true half siblings. Based on these results it would seem that SNPs are at a disadvantage over microsatellites. That being said, 19 of the super-jack's offspring were precocious males that returned in 2011. As these individuals are all of the same sex and return in a different year from their three-year old female siblings, the risk of inbreeding between these individuals is absent. The remaining 50 age-three offspring that made up the super-jack's contribution returned in 2012 and consisted of 9 males and 18 females

in one family and 9 males and 14 females in the other. In the context of a spawning matrix, the pairwise values one is concerned with are those that are generated between putative male and female half sibling crosses (9m x 18f and 9m x14f), which for this family make up a total of 188 possible pairs. Based on the proportion of half siblings that fell below the 0.125 cutoff point for all possible pairwise values generated from SNPs (29%), approximately 55 of the 188 pairs would fall below the 0.125 cutoff value. Although this is a relatively large number (compared to approximately 15 pairs in the microsatellite generated values using the 8% error rate for the super-jack's offspring), it is important to remember that the 170 three-year olds that returned in the 2012 spawn year were only distributed across 18 parent pairs. So few families comprised of numerous offspring would make it next to impossible to not mate related individuals, even if the mean pairwise SNP generated Rxy value for the super-jack's half-sibs were 0.25.

Overall, the results of this analysis indicate that SNPs perform comparably well to, if not better than, microsatellites at estimating relatedness between individuals under "normal" conditions. When the super-jack's offspring were removed from the half sibling group, the means of the distributions of known full- and half-sibs were nearly equal to (slightly larger) the expected means of the randomly generated distributions of full and half siblings. Additionally, with both marker types, reproductive success was significantly correlated to smaller values of Rxy in parent pairs, which has been shown to greatly increase the chance of survival in other hatchery populations of coho salmon (Conrad *et al.* 2013). Although the relatedness statistic, Rxy, is a useful proxy for

estimating kinship among individuals of unknown pedigree, this analysis also illuminates the level of variability that arises within the statistic when comparing two marker types in the same individuals, and ultimately supports the use of genetic tagging and pedigree inference to determine putative mates in populations that are facing extinction.

Conclusions and Future Directions

This study describes a large-scale genetic tagging effort by way of pedigree reconstruction in a hatchery population of coho salmon. I demonstrate that multi-locus SNP markers are powerful tools for generating accurate parentage-offspring assignments in a species that undergoes extensive migrations in both marine and freshwater environments. This knowledge will help us to understand the effects of hatchery practices on life history traits and fitness of hatchery salmon, as well as the potential negative effect these individuals might have on their wild counterparts. Ultimately this information will serve to improve hatchery practices and the overall fitness of hatchery populations, and in doing so, should help to conserve and protect natural populations of coho salmon.

References

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**(Suppl. 1):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**:4733-4746.

Aguilar, A., and J.C. Garza. 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.

Allen, R.L., T.K. Meekin, G.B. Pauley, and M.P. Fujihara. 1968. Mortality among Chinook salmon associated with the fungus *Dermocystidium*. Journal of the Fisheries Research Board of Canada **25**:2467-2475.

Allendorf, F.W., and G.H. Thorgaard. 1984. Tetraploidy and the Evolution of Salmonid Fishes in Evolutionary Genetics of Fishes, edited by B.J. Turner. Plenum Press, New York.

Anderson, E.C., and J.C. Garza. 2006. The power of single-nucleotide polymorphisms for large-scale parentage inference. Genetics **172**:2567-2582.

Anderson, E.C., R.S. Waples, and S.T. Kalinowski. 2008. An improved method for predicting the accuracy of genetic stock identification. Canadian Journal of Fisheries and Aquatic Sciences **65**:1475-1486.

Anderson, E.C. 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**:12p.

Appleby, A.E., J.M. Tipping, and P.R. Seidel. 2003. The effect of using two-year-old male coho salmon in hatchery broodstock on adult returns. North American Journal of Aquaculture **65**:60-62.

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**:100-103.

Banks, M.A., M.S. Blouin, M.A. Baldwin, V.K. Rashbrook, H.A. Fitzgerald, S.M. Blankenship, and D. Hedgecock. 1999. Isolation and inheritance of novel microsatellites in Chinook salmon (*Oncorhynchus tshawytscha*). Journal of Heredity **90**:281-288.

Banks, M.A., V.K. Rashbrook, M.J. Calavetta, C.A. Dean, and D. Hedgecock. 2000. Analysis of microsatellite DNA resolves genetic structure and diversity of Chinook salmon (*Oncorhynchus tshawytscha*) in Californias Central Valley. Canadian Journal of Fisheries and Aquatic Sciences **57**:915-927.

Bartholomew, J.L., and J.S. Foott. 2010. Compilation of information relating to Myxozoan disease effects to inform the klamath basin restoration agreement. Philosophical Transactions of the Royal Society of London, Biological Sciences **355**:1553-1562.

Beacham, T., R. Withler, and T. Stevens. 1996. Stock identification of Chinook salmon (*Oncorhynchus tshawytscha*) using minisatellite DNA variation. Canadian Journal of Fisheries and Aquatic Sciences **53**:380-394.

Beacham, T.D., and C.C. Wood. 1999. Application of microsatellite DNA variation to estimation of stock composition and escapement of Nass River sockeye salmon (*Oncorhynchus nerka*). Canadian Journal of Fisheries and Aquatic Sciences **56**:297-310.

Beacham, T.D., J.R. Candy, K.L. Jonsen, J. Supernault, M. Wetklo, L. Deng, K.M. Miller, R.E. Withler, and N. Varnavskaya. 2006. Estimation of stock composition and individual identification of Chinook salmon across the Pacific rim by use of microsatellite variation. Transactions of the American Fisheries Society **135**:861-888.

Beacham, T.D., M. Wetklo, L. Deng, and C. MacConnachie. 2011. Coho salmon population structure in North America determined from microsatellites. Transactions of the American Fisheries Society **140**:253-270.

Behnke, R.J. 2002. Trout and Salmon of North America. pp. 35-40. Chanticleer Press Inc. New York, NY.

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

Bell, E. and W.G. Duffy. 2007. Previously undocumented two-year old freshwater residency of juvenile coho salmon in Prairie Creek, California. Transactions of the American Fisheries Society **136**:966-970.

Berven, K.A., and T.A. Grudzien. 1990. Dispersal in the wood frog (*Rana sylvatica*): implications for genetic population structure. Evolution **44**:2047-2056.

Bilton, H.T., D.F. Alderdice, and J.T. Schnute. 1982. Influence of time and size at release of juvenile coho salmon (*Oncorhynchus kisutch*) on returns at maturity. Canadian Journal of Fisheries and Aquatic Sciences **39**:426-447.

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 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).

Blouin, M.S. 2003. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. TRENDS in Ecology and Evolution **18**:503-511.

Bouck, A., and T. Vision. 2007. The molecular ecologist's guide to expressed sequence tags. Molecular Ecology **16**:907-924.

Bradford, M.J., and J.R. Irvine. 2002. Land use, fishing, climate change, and the decline of Thompson River, British Columbia, coho salmon. Canadian Journal of Fisheries and Aquatic Sciences **57**:13-16

Brieuc, M.S.O, C.D. Waters, J.E. Seeb, and K.A. Naish. 2014. A dense linkage map for Chinook salmon (*Oncorhynchus tshawytscha*) reveals variable chromosomal divergence after an ancestral whole genome duplication event. G3:Genes Genomes Genetics **4**:447-460

Brown, L. R., P.B. Moyle, and R.M. Yoshiyama. 1994. Historical decline and current status of coho salmon in California. North American Journal of Fisheries Management **1**:237-261

Brumfield, R.T., P. Beerli, D.A. Nickerson, and S.V. Edwards. 2003. The utility of single nucleotide polymorphisms in inferences of population history. Trends in Ecology & Evolution **18**:249-256.

Buckland, K.A., M. Manks, and D. Hedgecock. 2007. Assessing genetic diversity of protected coho salmon (*Oncorhynchus kisutch*) populations in California. Canadian Journal of Fisheries and Aquatic Sciences **64**:30-42

Buhle, E.R., K.K. Holsman, M.D. Scheuerell, and A. Albaugh. 2009. Using an unplanned experiment to evaluate the effects of hatcheries and environmental variation on threatened populations of wild salmon. Biological Conservation **142**:2449-2455.

California Hatchery and Genetic Management Plan Version 10. 2013. For Iron Gate Hatchery Coho Salmon. Prepared for the National Oceanic and Amospheric Administration, National Marine Fisheries Service.

California Hatchery Scientific Review Group. 2012. Iron Gate Hatchery Coho Program Report. California Hatchery Review Project, Apendix VIII. Prepared for the U.S. Fish and Wildlife Service and Pacific States Marine Fisheries Commission.

Campbell, N.R., and S.R. Narum. 2011. Development of 54 novel single-nucleotide polymorhism (SNP) assays for sockeye and coho salmon and assessment of available SNPs to differentiate stocks within the Columbia River. Molecular Ecology Resources **11**(Suppl. 1):20-30.

Carlson, S.M., and T.R. Seamons. 2008. A review of quantitative genetic components of fitness in salmonids: implications for adaptation to future change. Evolutionary Applications **1**:222-238.

Chesney, D. 2007. Recovery of fall-run Chinook and coho salmon at Iron Gate Hatchery, 2006. California Department of Fish and Game, Klamath River Project, Final Report.

Chesney, D. 2009. Recovery of fall-run Chinook and coho salmon at Iron Gate Hatchery September 24, 2010 to December 15, 2010. California Department of Fish and Game, Klamath River Project, Final Report.

Chesney, D., and M. Knechtle. 2011. Recovery of fall-run Chinook and coho salmon at Iron Gate Hatchery, 2006. California Department of Fish and Game, Klamath River Project, Final Report.

Christie, M.R., M.L. Marine, R.A. French, and M.S. Blouin. 2012. Genetic adaptation to captivity can occur in a single generation. Proceedings of the National Academy of Sciences **109**:238-242.

Christie, M.R., R.A. French, M.L. Marine, and M.S. Blouin. 2013. How much does inbreeding contribute to the reduced fitness of hatchery-born Steelhead (*Oncorhynchus mykiss*) in the wild? Journal of Heredity **105**:111-119.

Clark, A.G., M.J. Hubisz, C.D. Bustamante, S.H. Williamson, and R. Nielsen. 2005. Ascertainment bias in studies of human genome-wide polymorphism. Genome Research **15**:1496-1502.

Clemento, A.J., E.C. Anderson, D. Boughton, D. Girman, and J.C. Garza. 2009. Population genetic structure and ancestry of *Oncorhynchus mykiss* populations above and below dams in south-central California. Conservation Genetics **10**:1321-1336.

Clemento, A.J., A. Abadía-Cardoso, H.A. Starks, and J.C. Garza. 2011. Discovery and characterization of single nucleotide polymorphisms in Chinook salmon, *Oncorhynchus tshawytscha*. Molecular Ecology Resources **11**(Suppl. 1):50-66.

Cole, J. 2000. Coastal sea surface temperature and coho salmon production off the northwest United States. Fisheries Oceanography **9**:1-16.

Condrey, M.J., and P. Bentzen. 1998. Characterization of coastal cutthroat trout (*Oncorhynchus clarki clarki*) microsatellites and their conservation in other salmonids. Molecular Ecology **59**:209-215.

Conrad, J.L., E.A. Gilbert-Horvath, and J.C. Garza. 2013. Genetic and phenotypic effects on reproductive outcomes for captively-reared coho salmon, *Oncorhynchus kisutch*. Aquaculture **404-405**:95-104.

Cook, R. C. 1982. Stock identification of sockeye salmon (*Oncorhynchus nerka*) with scale pattern recognition. Canadian Journal of Fisheries and Aquatic Sciences **39**:611-617.

Crozier, W.W., and G.J.A. Kennedy. 2002. Impact of coded wire tags on marine survival of wild Atlantic salmon (*Salmo salar* L.) migrating from the R. Bush, Northern Ireland. Fisheries Research **59**:209-215.

de Fromentel, C.C., F. Pakdel, A. Chapus, C. Baney, P. May, and T. Soussi. 1992. Rainbow trout p53: cDNA cloning and biochemical characterization. Genetics **112**:241-245.

Department of Water Resources. 2004. The effects of the Feather River Hatchery on naturally spawning salmonids. SP-F9 Final Report. Available from: http://orovillerelicensing.water.ca.gov/wg-reports_envir.html

DeWoody, J.A. 2005. Molecular approaches to the study of parentage, relatedness and fitness: practical applications for wild animals. Journal of Wildlife Management **69**:1400-1418.

Drucker, B. 1972. Some life history characteristics of coho salmon of the Karluk River system, Kodiak Island, Alaska. Fishery Bulletin **70**:79-94.

Du S.J., R.H. Devlin, and C.L. Hew. 1993. Genomic structure of Growth Hormone Genes in Chinook salmon (*Oncorhynchus tshawtscha*):presence of two functional genes, GH-I and GH-II, and a male-specific pseudogene, GHpsi. DNA and Cell Biology **12**:739-751.

Elliott, D.G., and R.J. Pascho. 2001. Evidence that coded-wire-tagging procedures can enhance transmission of *Renibacterium salmoninarum* in Chinook salmon. Journal of Aquatic Animal Health **13**:181-193.

Federal Register. 1995. Endangered and Threatened Species; Proposed Threatened Status for Three Contiguous ESUs of Coho Salmon Ranging from Oregon Through Central California. Federal Register **60**: 38011-38030. Office of the Federal Register, National Archives and Records Administration (NARA), College Park, MD.

Federal Register. 1997. Endangered and Threatened Species; Threatened Status for Southern Oregon/Northern California Coast Evolutionarily Significant Unit (ESU) of Coho Salmon. Federal Register **62**:24588-24609. Office of the Federal Register, NARA, College Park, MD.

Federal Register. 2005. Endangered and Threatened Species: Final Listing Determinations for 16 ESUs of West Coast Salmon, and Final 4(d) Protective Regulations for Threatened Salmonid ESUs. Federal Register **70**:37160-37204. Office of the Federal Register, NARA, College Park, MD.

Flannery, B.G., J.K. Wenburg, and A.J. Gharrett. 2007. Variation of amplified fragment length polymorphisms in Yukon River chum salmon: population structure and application to mixed-stock analysis. Transactions of the American Fisheries Society **136**:911-925.

Ford, M.J., H. Fuss, B. Boelts, E. LaHood, J. Hard, and J. Miller. 2006. Changes in run timing and natural smolt production in a naturally spawning coho salmon (*Oncorhynchus kisutch*) population after 60 years of intensive hatchery supplementation. Canadian Journal of Fisheries and Aquatic Sciences **63**:2343-2355.

Fowler, E.V., B.A. Houlden, W.B. Sherwin, P. Hoeben, and P. Timms. 1998. Genetic variation in captive koalas (*Phascolarctos cinereus*): parentage determination and individual identification. Biochemical Genetics **36**:193-206.

Frissell, C.A. 1993. Topoplogy of extinction and endangerment of native fishes in the Pacific Northwest and California (U.S.A.). Conservation Biology **7**:342-354.

Funk, W.C., J.A. Tyburczy, K.L. Knudsen, K.R. Lindner, and F.W. Allendorf. 2005. Genetic basis of variation in morphological and life-history traitsof a wild population of pink salmon. Journal of Heredity **96**:24-31.

Garza, J.C., and E.G. Williamson. 2001. Detection of reduction in population size using data from microsatellite loci. Molecular Ecology **10**:305-318.

Garza, J.C., and E. Gilber-Horvath. 2003. Report on the Genetics of Coho Salmon (*Oncorhynchus kisutch*) held at Warm Springs (Don Clausen) Hatchery for Recovery Efforts in the Russian River. December 2003

Garza, J.C., and E.C. Anderson. 2007. Large scale parentage inference as an alternative to coded-wire tags for salmon fishery management. In: PSC Genetic Stock Identification Workshop (May and September 2007): Logistics Workgroup final report and recommendations, p. 48-55. Pacific Salmon Commission, Vancouver, British Columbia, Canada. Available from: http://www.psc.org/info_genetic_stock_id.htm

Good, T.P., R.S. Waples, and P. Adams. 2005. Updated status of federally listed ESUs of West Coast salmon and steelhead, p.598. NOAA Technical Memorandum (NMFS-NWFSC-66).

Gross, M.R. 1991. Salmon breeding behavior and life history evolution in changing environments. Ecology **72**(4):1180-1186.

Groot, C., and L. Margolis. 1991. *Pacific Salmon Life Histories.* Vancouver: UBC Press.

Habicht, C., S. Sharr, D. Evans, and J.E. Seeb. 1998. Coded wire tag placement affects homing ability of pink salmon. Transactions of the American Fisheries Society **127**:652-657.

Hager, R.C. and R.E. Noble. Relation of size at release of hatchery-reared coho salmon to age, size, and sex composition of returning adults. The Progressive Fish-Culturist **38**:144-147.

Hamilton, J.B., G.L. Curtis, S.M. Snedaker, and D.K. White. 2005. Distribution of anadromous fishes in the Upper Klamath River watershed prior to hydropower dams - a synthesis of the historical evidence. Fisheries **30**:10-20.

Hampton, M. 2005. Recovery of Chinook and coho salmon at Iron Gate Hatchery October 4, 2004 to December 20, 2004. California Department of Fish and Game, Klamath River Project, Final Report.

Hankin, D.G., J.W. Nicholas, and T.W. Downey. 1993. Evidence for inheritance of age of maturity in Chinook sa *Oncorhynchus tshawytscha.* Canadian Journal of Fisheries and Aquatic Sciences **50**:347-358.

Hankin, D.G., J.H. Clark, R.B. Deriso, J.C. Garza, G.S. Morishima, B.E. Riddell, C. Schwarz, and J.B. Scott. 2005. Report of the expert panel on the future of the coded wire tag recovery program for Pacific salmon. Pacific Salmon Commission Technical Report **18**: 230p. Available from: http://www.psc.org/pubs/psctr18.pdf

Hare, S.R., N.J. Mantua, and R.C. Francis. 1999. Inverse production regimes: Alaska and West Coast Pacific Salmon. Fisheries **24**:6-14.

Hallett, S.L., R.A. Ray, C.N. Hurst, R.A. Holt, G.R. Buckles, S.D. Atkinson, and J.L. Bartholomew. 2012. Density of the waterborne parasite *Ceratomyxa shasta* and Its biological effects on salmon. Applied and Environmental Microbiology **78**:3724-288.

Håstein, T. and T. Lindstad. 1991. Diseases in wild and cultured salmon: possible interaction. Aquaculture **98**:277-3731.

Heath, D.D., C.W. Fox, and J.W. Heath. 1999. Maternal effects on offspring size: variation through early development of Chinook salmon. Evolution **53**:1605-1611.

Herrera, E.A. 1992. Growth and dispersal of capybaras (*Hydrochaeris* hydrochaeris*) in the Llanos of Venezuela. The Zoological Society of London **228**:307-316.

Hess, J.E., A.P. Matala, and S.R., Narum. 2011. Comparison of SNPs and microsatellites for fine-scale application of genetic stock identification of Chinook salmon in the Columbia River basin. Molecular Ecology Resources **11**(Suppl. 1):137-149.

Hilborn, R. 1992. Hatcheries and the future of salmon in the Northwest. Fisheries **17**: 5-8.

Hoffman, J.I., P.N. Trathan, and W. Amos. 2006. Genetic tagging reveals extreme site fidelity in territorial male Antarctic fur seals *Arctocephalus gazella*. Molecular Ecology **15**: 3841-3847.

Holtby, L.B., B.C. Andersen, and R.K. Kadowaki. 1990. Importance of smolt size and early ocean growth to interannual variability in marine survival of coho salmon (*Oncorhynchus kisutch*). Canadian Journal of Fisheries and Aquatic Sciences **47**:2181-2194.

Iwamoto, R.N., B.A. Alexander, and W.K. Hershberger. 1984. Genotypic and environmental effects of the incidence of sexual precocity in coho salmon (*Oncorhynchus kisutch*). Aquaculture **43**:105-121.

Johnston, J.P., W.J. Peach,R.D. Gregory, and S.A. White. 1997. Survival of tropical and temperate passerines: a trinidadian perspective. The American Naturalist **150**:771-789.

Konovalov, D.A., C. Manning, and M.T. Henshaw. 2004. KINGROUP: a program for pedigree relationship reconstruction and kin group assignments using genetic markers. Molecular Ecology Notes **4**:779-782.

Landry, C., D. Garant, P. Duchesne, and L. Bernatchez. 2001. Good genes as heterozygosity: the major histocompatibility complex and mate choice in Atlantic salmon (*Salmo salar*). Proceedings of the Royal Society Biological Sciences **268**:1279-1285.

Lestelle, L.C. 2007. Coho Salmon (*Oncorhynchus kisutch*) Life History Patterns in the Pacific Northwest and California. Final Report, U.S. Bureau of Reclamation, Klamath Office Area.

McConnell, S.K., P. O'Reilly, L. Hamilton, J.M. Wright, and P. Bentzen. 1995. 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, J.A., and C.A. Simenstad. 1997. SA comparative assessment of a natural and created estuarine slough as rearing habitat for juvenile Chinook and coho salmon. Estuaries **20**:792-806.

Milner, G.B., D.J. Teel, F.M. Utter, and G.A. Winans. 1985. A genetic method of stock identification in mixed populations of Pacific salmon, *Oncorhynchus spp.* Marine Fisheries Review **47**:1-8.

Morin, P., G. Luikart, and R. Wayne. 2004. SNPs in ecology, evolution and conservation. Trends in Ecology and Evolution **19**:208-216.

Morrison, J., and D. Zajac. 1987. Histologic effect of coded wire tagging in chum salmon. North American Journal of Fisheries Management **7**:439-441.

Morton, M.L. 1992. Effects of sex and birth date on premigration biology, migration schedules, return dates and natal dispersal in the mountain white-crowned sparrow. The Cooper Ornithological Society **94**:117-133.

Narum, S.R., M.A. Banks, T.D. Beacham, M.R. Bellinger, M.R. Campbell, J. Dekoning, A. Elz, C.M. Guthrie, C. Kozfkay, K.M. Miller, P. Moran, R. Phillips, L.W. Seeb, C.T. Smith, K. Warheit, S.F. Young, and J.C. Garza. 2008. Differentiating salmon populations at broad and fine geographical scales with microsatellites and single nucleotide polymorphisms. Molecular Ecology **17**:3464-3477.

Nehlsen, W., J.E. Williams, and J.A. Lichatowich. 1991. Pacific salmon at the crossroads: stocks at risk from California, Oregon, Idaho, and Washington. Fisheries **16**:4-21.

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

Nielsen, R. 2005. Molecular signatures of natural selection. Annual Revue of Genetics **39**:197-218.

O'Reilly, P.T., L.C. Hamilton, S.K. McConnell, and J.M. Wright. 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.

Palsbøll, P.J., J. Allen, M. Bérubé, P.J. Clapham, T.P. Feddersen, P.S. Hammond, R.R. Hudson, H. Jørgensen, S. Katona, A.H. Larsen, F. Larsen, J. Lien, D.K. Mattila, J. Sigurjónsson, R. Sears, T. Smith, R. Sponer, P. Stevick, and N. Oien. 1997. Genetic tagging of humpback whales. Nature **388**:767-769.

Park, S.D.E. 2001. Trypanotolerance in West African cattle and the population genetic effects of selection. Ph.D. thesis: University of Dublin

Peakall, R., D. Ebert, R. Cunningham, and D. Lindenmayer. 2006. Mark-recapture by genetic tagging reveals restricted movements by bush rats (*Rattus fuscipes*) in a fragmented landscape. Journal of Zoology **268**:207-216.

Pearse, D.E., C.J. Donohoe, and J.C. Garza. 2007. Population genetics of steelhead (*Oncorhynchus mykiss*) in the Klamath River. Environmental Biology of Fishes **80**:377-387.

Pearse, D.E. and E.C. Anderson. 2009. Multiple paternity increases effective population size. Molecular Ecology **18**:3124-3127.

Picoult-Newberg, L., T.E. Ideker, M.G. Pohl, S.L. Taylor, M.A. Donaldson, D.A. Nickerson, and M. Boyce-Jacino. 1999. Mining SNPs from EST databases. Genome Research **9**:167-174.

Queller, D.C., and K.F. Goodnight. 1989. Estimating relatedness using genetic markers. Evolution **43**:258-275.

Quinn, T.P., and C.A. Busack. 1985. Chemosensory recognition of siblings in juvenile coho salmon (*Oncorhynchus kisutch*). Animal Behaviour **33**:51-56.

Quinn, T.P., and N.P. Peterson. 1996. The influence of habitat complexity and fish size on over-winter survival and growth of individually marked juvenile coho salmon (*Oncorhynchus kisutch*) in Big Beef Creek, Washington. Canadian Journal of Fisheries and Aquatic Sciences **53**:1555-1564.

Quiñones, R.M., M.L. Johnson, and P.B. Moyle. 2014. Hatchery practices may result in replacement of wild salmonids: adult trends in the Klamath basin, California. Environmental Biology of Fishes **97**:233-246.

Reimchen, T.E., and N.F. Temple. 2004. Hydrodynamic and phylogenetic aspects of the adipose fin in fishes. Canadian Journal of Zoology **82**:910-916.

Rexroad, C.E., R.L. Coleman, W.K. Hershberger, and J. Killefer. 2002. Rapid communication: Thirty-eight polymorphic microsatellite markers for mapping in rainbow trout. Journal of Animal Science **80**:541-542.

Richey, J. 2006. Recovery of fall-run Chinook and coho salmon at Iron Gate Hatchery, 2005. California Department of Fish and Game, Klamath River Project, Final Report.

Ricker, W.E. 1981. Changes in the average size and average age of Pacific Salmon. Canadian Journal of Fisheries and Aquatic Sciences **38**:1636-1656.

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., and H.J. Skaletsky. 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. Software available from: http://fokker.wi.mit.edu/primer3/input.htm

Ruckelshaus, M.H., P. Levin, J.B. Johnson, and P.M. Kareiva. 2002. The Pacific Salmon Wars: What science brings to the challenge of recovering species. Annual Review of Ecology and Systematics **33**:665-706.

Sánchez, C.C., T.P.L. Smith, R.T. Wiedmann, R.L. Vallejo, M. Salem, J. Yao, and C.E. Rexroad. 2009. Single nucleotide polymorphism disovery in rainbow trout by deep sequencing of a reduced representation library. BCM Genomics **10**:559-567.

Sandercock, F.K. 1991. Life history of coho salmon (*Onchorhynchus kisutch*). In: Pacific Salmon Life Histories (eds C. Groot and L. Margolis). pp. 395-445. University of British Columbia Press, Vancouver.

Sawaya, M.A., S.T. Kalinowski, and A.P. Clevenger. 2014. Genetic connectivity for two bear species at wildlife crossing structures in Banff National Park. Proceedings of the Royal Society of Biological Sciences **281**:1-10.

Scribner, K.T., J.R. Gust, and R.L. Fields. 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.

Seeb, L.W., and J.E. Seeb. 2006. Report of the second annual SNP workshop, Anchorage, Alaska, USA. North Pacific Anadromous Fish Commission, NPAFC Doc 996, p.11.

Seeb, L.W., A. Antonovich, M.A. Banks, T.D. Beacham, M.R. Bellinger, S.M. Blankenship, M.R. Campbell, N.A. Decovich, J.C. Garza, C.M. Guthrie III, T.A. Lundrigan, P. Moran, S.R. Narum, J.J. Stephenson, K.J. Supernault, D.J. Teel, W.D. Templin, J.K. Wenburg, S.F. Young, and C.T. Smith. 2007. Development of a standardized DNA database for Chinook salmon. Fisheries **32**:540-552.

Shapovalov, L. and A.C. Taft. 1954. The life histories of the steelhead rainbow trout (*Salmo gairdneri gairdnerii*) and silver salmon (*Oncorhynchus kisutch*) with special reference to Waddel Creek, California, and recommendations regarding their management. California Department of Fish and Game, Fisheries Bulletin No. 98.

Small, M.P., R.E. Withler, and T.D. Beacham. 1998. Population structure of British Columbia coho salmon, *Oncorhynchus kisutch*, based on microsatellite DNA variation. Fisheries Bulletin **96**:843-858.

Smith, C.T., B.F. Koop, and R.J. Nelson. 1998. Isolation and characterization of coho salmon (*Oncorhynchus kisutch*) microsatellites and their use in other salmonids. Molecular Ecology **7**:1614-1616.

Smith, C.T., R.J. Nelson, C.C. Wood, and B.F. Koop. 2001. Glacial biogeography of North American coho salmon (*Oncorhynchus kisutch*). Molecular Ecology **10**:2775-2785

Smith, C.T., C.M. Elfstrom, L.W. Seeb, and J.E. Seeb. 2005. Use of sequence data from rainbow trout and Atlantic salmon for SNP detection in Pacific salmon. Molecular Ecology **14**, 4193-4203.

Smith, C.T., L. Park, D. VanDoornik, L.W. Seeb, and J.E. Seeb. 2006. Characterization of 19 single nucleotide polymorphism markers for coho salmon. Molecular Ecology Notes **6**:715-720.

Steele, C.A., E.C. Anderson, M.W. Ackerman, M.A. Hess, N.R. Campbell, S.R. Narum, and M.R. Campbell. 2013. A validation of parentage-based tagging using hatchery steelhead in the Snake River basin. Canadian Journal of Fisheries and Aquatic Sciences **70**:1046-1054.

Sturm, E.A., E.A. Gilbert-Horvath, J.C. Garza, and R.B. MacFarlane. 2009. Creation of a captive broodstock program for southern coho salmon (*Oncorhynchus kisutch*): results from the initial rearing and spawning of the first brood year. NOAA Technical Memorandum NMFS, NOAA-TM-NMFS-SWFSC-436.

Sutcliffe, O.L., and C.D. Thomas. 1997 Area-dependent migration by ringlet butterflies generates a mixture of patchy population and metapopulation attributes. Oecologia **109**:229-234.

Taylor, E.B. 1991. A review of local adaptation in Salmonidae, with particular reference to Pacific and Atlantic salmon. Aquaculture **98**:185-207.

Templin, W.D., J.E. Seeb, J.R. Jasper, A.W. Barclay, and L.W. Seeb. 2011. Genetic differentiation of Alaska Chinook salmon: the missing link for migratory studies. Molecular Ecology Resources **11**(Suppl. 1):226-246.

Thrower, F.P., and J.J. Hard. 2009. Effects of a single event of close inbreeding on growth and survival in steelhead. Conservation Genetics **10**:1299-1307.

Trajano, E. 1997. Population ecology of *Trichomycterus itacarambiensis*, a cave catfish from eastern Brazil (Siluriformes, Trichomycteridae). Environmental Biology of Fishes **50**:357-369.

Unger, S.D., N.G. Burgmeier, and R.N. Williams. 2012. Genetic markers reveal high PIT tag retention rates in giant salamanders (*Cryptobranchus alleganiensis*). Amphibia-Reptilia **33**:313-317.

Vignal, A., D. Milan, M. SanCristobal, and A. Eggen. 2002. A review on SNP and other types of molecular markers and their use in animal genetics. Genetics, Selection, Evolution **34**:275-305.

Wang, S., J.J. Hard, and F. Utter. 2002. Salmonid inbreeding: a review. Reviews in Fish Biology and Fisheries **11**:301-319.

Wang, C.M., L.C. Lo, Z.Y. Zhu, G. Lin, F. Feng, J. Li, W.T. Yang, J. Tan, R. Chou, H.S. Lim, L. Orban, and G.H. Yue. 2008. Estimating reproductive success of brooders and heritability of growth traits in Asian sea bass (*Lates calcarifer*) using microsatellites. Aquaculture Research **39**:1612-1619.

Waples, R.S. 1991. Genetic interactions between hatchery and wild salmonids: lessons from the Pacific Northwest. Canadian Journal of Fisheries and Aquatic Sciences **48**(Suppl. 1):124-133.

Weir, B.S., and C.C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution **38**:358-1370.

Weitkamp, L.A., T.C. Wainwright, G.J. Bryant, G.B. Milner, D.J. Teel, R.G. Kope, and R.S. Waples. 1995. Status review of coho salmon from Washington, Oregon, and California. NOAA Technical Memorandum, NMFS-NWFSC-24.

Weslien, J., and A. Lindelow. 1990. Recapture of marked spruce bark beetles (*Ips typographus*) in pheromone traps using area-wide mass trapping. Canadian Journal of Forest Restoration **20**:1786-1790.

Williamson, K.S., J.F. Cordes, and B. May. Characterization of microsatellite loci in chinook salmon (*Oncorhynchus tshawytscha*) and cross-species amplification in other salmonids. Melecular Ecology Resources **2**:17-19.

Woods, J.G., D. Paetkau, D. Lewis, B.N. McLellan, M. Proctor, and C. Strobeck. 1999. Genetic tagging of free-ranging black and brown bears. Wildlife Society Bulletin **27**:616-627.

Young, K.A. 1999. Environmental correlates of male life history variation among coho salmon populations from two Oregon coastal basins. Transactions of the American Fisheries Society **128**:1-16.