EVALUATION OF PARAMETER ESTIMATION AND FIELD APPLICATION OF TRANSGENERATIONAL GENETIC MARK-RECAPTURE

By

Ryan William Whitmore

A Thesis Presented to

The Faculty of Humboldt State University

In Partial Fulfillment of the Requirements for the Degree

Master of Science in Natural Resources: Fisheries

Committee Membership

Dr. Andrew Kinziger, Committee Chair

Dr. David Hankin, Committee Member

Dr. Darren Ward, Committee Member

Dr. Alison O'Dowd, Graduate Coordinator

December 2016

ABSTRACT

EVALUATION OF PARAMETER ESTIMATION AND FIELD APPLICATION OF TRANSGENERATIONAL GENETIC MARK-RECAPTURE

Ryan William Whitmore

Use of a genetic-based analogue of the traditional mark-recapture method (transgenerational genetic mark-recapture, tGMR) is rapidly expanding as a means to estimate total escapement of Pacific salmon. The tGMR approach is similar to the simple Lincoln-Peterson mark-recapture method. In tGMR, adults returning to fresh water to spawn are collected in the first sampling occasion and their juvenile offspring that are migrating out to sea are collected during the second sampling occasion. Recaptures are determined by the number of parent-offspring pairs identified through genetic parentage analysis of the adult and juvenile collections. Two versions of tGMR are currently in use, referred to as the "with replacement" and the "without replacement" models. For each version, parentage analysis is used to estimate model parameters. I evaluated accuracy of tGMR parameters estimated by genetic parentage analysis by conducting a series of simulations that mimicked application of the approach for estimating escapement of a small northern California coho salmon population. Accuracy was evaluated by comparing known values of the parameters taken from the simulated pedigrees to estimated values based upon parentage analysis of SNP genotypes using the software COLONY. All parentage-based parameter estimates were biased, (ranging from -0.40 to 0.23) indicating

improvements in parentage analysis are needed for applications of tGMR. To further evaluate tGMR, I applied this method to coho salmon in two northern California streams resulting in total escapement estimates using the "with replacement" and "without replacement" models of 576 and 444 (Mill Creek, 2011-2012), 131 and 193 (Mill Creek, 2012-2013), and 430 and 468 (Freshwater Creek, 2012-2013). The tGMR approach shows promise for highly fecund species because the number of individuals captured during the second sampling occasion can greatly exceed the adult population size. This can possibly lead to lower variance in tGMR estimates in comparison to traditional mark-recapture estimators, but improvements in genetic parentage analysis are needed to reduce or eliminate bias from parentage analysis that results in biased estimates of total escapement using tGMR.

ACKNOWLEDGEMENTS

Funding for this project was provided by California Department of Fish and Wildlife's Fisheries Restoration Grant Program. Additional funding was provided through scholarships by the Marin Rod and Gun Club, Danielle Plum Zumbrun Memorial Scholarship, and the Fresno Area Fly Fisheries Club. I want to thank Dr. Andrew Kinziger for his endless patience and knowledge that allowed for this project to be possible. Special thanks to Dr. David Hankin for his analytical guidance and comments, and to Dr. Darren Ward for his invaluable comments and discussion. Much appreciation to Seth Ricker, Justin Garwood, Colin Anderson, and the staff at California Department of Fish and Wildlife, Arcata, for providing tissue samples, and to Justin Garwood for the beautiful locator map. I would also like to thank Dr. Carlos Garza and the MEGA team staff at NOAA's Southwest Fisheries Science Center for processing all the tissue samples and to Dr. Eric Anderson (MEGA team member) for his vast knowledge of parentage analysis and enthusiasm during our discussions. A very special thanks to Dr. Michael Miller at UC Davis for his use of the FARMII cluster. Without his help, the simulation portion of the project would not have been possible. Finally, and most importantly, I thank my mom for her constant support throughout this very long journey. Without her, my future in fisheries management would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: EVALUATION OF THE ACCURACY OF GENETIC PARENTAGE RECONSTRUCTIONS USED IN TRANSGENERATIONAL GENETIC MARK-RECAPTURE

CHAPTER 1

EVALUATION OF THE ACCURACY OF GENETIC PARENTAGE RECONSTRUCTIONS USED IN TRANSGENERATIONAL GENETIC MARK-RECAPTURE

ABSTRACT

Transgenerational genetic mark-recapture (tGMR) is rapidly expanding in popularity as a tool for estimating salmonid total escapement. However, there has been no evaluation of the accuracy of genetic parentage reconstructions that are needed to implement this method. The tGMR approach is a modification of the simple Lincoln-Peterson mark-recapture method. In tGMR, spawning adults are collected in the first sampling occasion and their offspring are collected at outmigration during the second sampling occasion. Recaptures are the number of parent-offspring pairs identified through genetic parentage analysis of the adult and juvenile collections. Two versions of tGMR are currently in use, referred to as the "with replacement" and the "without replacement" models. In the "with replacement" model, genetic parentage analysis is used to estimate the total number of times adults in the first sampling occasion (*M*) are assigned as parents of individuals from the juvenile sample, or the number of recaptures (*R*). In the "without replacement" model, parentage analysis is used to determine (i) the total number of parents required to produce the juvenile sample (*C*'), and (ii) the number of unique parents from the first sampling occasion (*M*) that are assigned as parents of individuals from the juvenile genotypes in the second sampling occasion (*R*'). To evaluate accuracy of genetic parentage analysis for estimating *R*, *C*', and *R*', I simulated an age-structured population of coho salmon (escapement= 800, offspring= 14,500), including pedigrees and individual multilocus genotypes. Simple random samples of adults and their offspring were selected to represent the first and second sampling

occasions as in tGMR. Accuracy was evaluated by comparing known values of *R*, *C*', and *R*' taken from the simulated pedigrees to estimated values of these parameters based upon parentage analysis of SNP genotypes using the COLONY software. Simulations were run for 32 combinations of loci (93 SNPs and 186 SNPs), parent subsample size (*M*= 40 and 160), and juvenile sample sizes (*n*= 500, 1000, 2000, 3000, 4000, 6000, 8000, and 10000). Simulation results indicated proportional bias ranged from 0.05 to -0.06 for *R*, from -0.40 to 0.23 for *C*', and from 0.03 to 0.00 for *R*'. My analysis suggests improvements in genetic parentage analysis, potentially resulting from more powerful molecular marker sets (e.g., more loci or multiallelic loci) will be needed to eliminate bias resulting from parentage analysis for tGMR. I recommend simulations as a tool for evaluating the extent to which genetic parentage analysis may bias tGMR estimates of salmonid escapement.

INTRODUCTION

Advancements in genetic technology have increased the amount of genetic information that can be generated from wild populations of organisms (Blouin 2003). This has resulted in the application of genetic parentage analysis as a "tagging" approach in wild populations. Parentage-based tagged methods have been useful for estimating trait heritability, age structure, reproductive success, and abundance of wild populations (Jones and Avise 1997; Pearse et al. 2001; Fiumera et al. 2002; Abadía-Cardoso et al. 2013; Bravington et al. 2016). The implementation of maximum likelihood methods has increased precision of parentage analysis (Wang 2004; Wang 2012; Wang 2013; Wang and Santure 2009). Further, the development of user-friendly programs such as FRANz (Reister et al. 2009), COLONY (Jones and Wang 2009), and SNPPIT (Anderson 2010) have made parentage analysis much more accessible. Researchers have applied parentage analysis to wild populations to estimate the number of breeders (Israel and May 2010), but also the total number of individuals present at a given time (Rawding et al. 2014). One application of such a model was used to estimate the total escapement of Chinook salmon (*Oncorhynchus tshawytscha*), and coined transgenerational genetic mark-recapture (tGMR) (Rawding et al. 2014).

The tGMR approach is a modification of the Lincoln-Peterson mark-recapture method. The Lincoln-Peterson estimator is a two-occasion model (Seber 1982):

$$\widehat{N} = \frac{MC}{R}$$

where the population size ($\widehat{N}$) is estimated by marking (*M*) a fraction of the population in the first sampling occasion and then capturing (*C*) individuals during a second sampling occasion. The number of marked individuals present in *C* during the second sampling occasion are the recaptures, *R*. Following model assumptions, the expected fraction of marked individuals on the second sampling occasion (R/C) should equal the fraction of marked individuals established after the first sampling occasion (M/N) (Williams et al. 2002).

Transgenerational genetic mark-recapture takes advantage of high throughput genotyping technologies and advances in genetic parentage algorithms. In tGMR, the first sampling occasion occurs when tissue is collected from live adults as they enter the river or from carcasses encountered during spawning surveys. The second sampling occasion occurs when tissue is collected from juveniles in traps during the out-migrant season. There are currently two versions of the tGMR method, termed the "with replacement" and "without replacement" models (Rawding et al. 2014).

In the binomial or "with replacement" method, total escapement is estimated as:

$$\widehat{N}_{wr} = \frac{MC}{R}$$

where adult salmon sampled in the first occasion and successfully genotyped are considered the marks (*M*). The recapture sample size, *C*, is equal to twice the number of juvenile out-migrants sampled in the second occasion (*n*) and successfully genotyped (*C*= 2*n*). Each genotyped juvenile has the potential of being assigned to both a male and female parent in *M* through parentage analysis and therefore represents two possible

recapture opportunities. A juvenile from the second sample could be assigned to no parents in *M*, one parent in *M* or two parents in *M*. The sum of the total parentage assignments (0, 1 or 2) for each juvenile are considered the recaptures (*R*) for the "with replacement" model. This is referred to as the sum of the number of times adults present in *M* are detected in the juvenile sample and the accuracy of this value is dependent upon the performance of genetic parentage analysis.

The hypergeometric approach was also presented in Rawding et al. (2014) and applies a "without replacement" sampling framework, which is also appropriate for the estimation of abundance (Seber 1982):

$$\widehat{N}_{wo} = \frac{MC'}{R'}$$

In comparison to the binomial method, the hypergeometric approach requires an estimate of the total number of distinct parents, *C*', that gave rise to the juvenile subsample of size *n*. This is the total number of individuals from *M* that were assigned as parents of juvenile through genetic parentage analysis plus the number of additional "unmarked" parents inferred to have been present by parentage analysis. The recaptures, *R'*, are the number of distinct parents from the first sampling occasion (*M*) that are assigned as parents of individuals in *n*.

Transgenerational genetic mark-recapture has been used to estimate total escapement in multiple populations of Chinook salmon in the Pacific Northwest (Rawding et al. 2014; Seamons et al. 2014; Seamons et al. 2015). However, while escapement estimates using tGMR have been compared to estimates derived from more

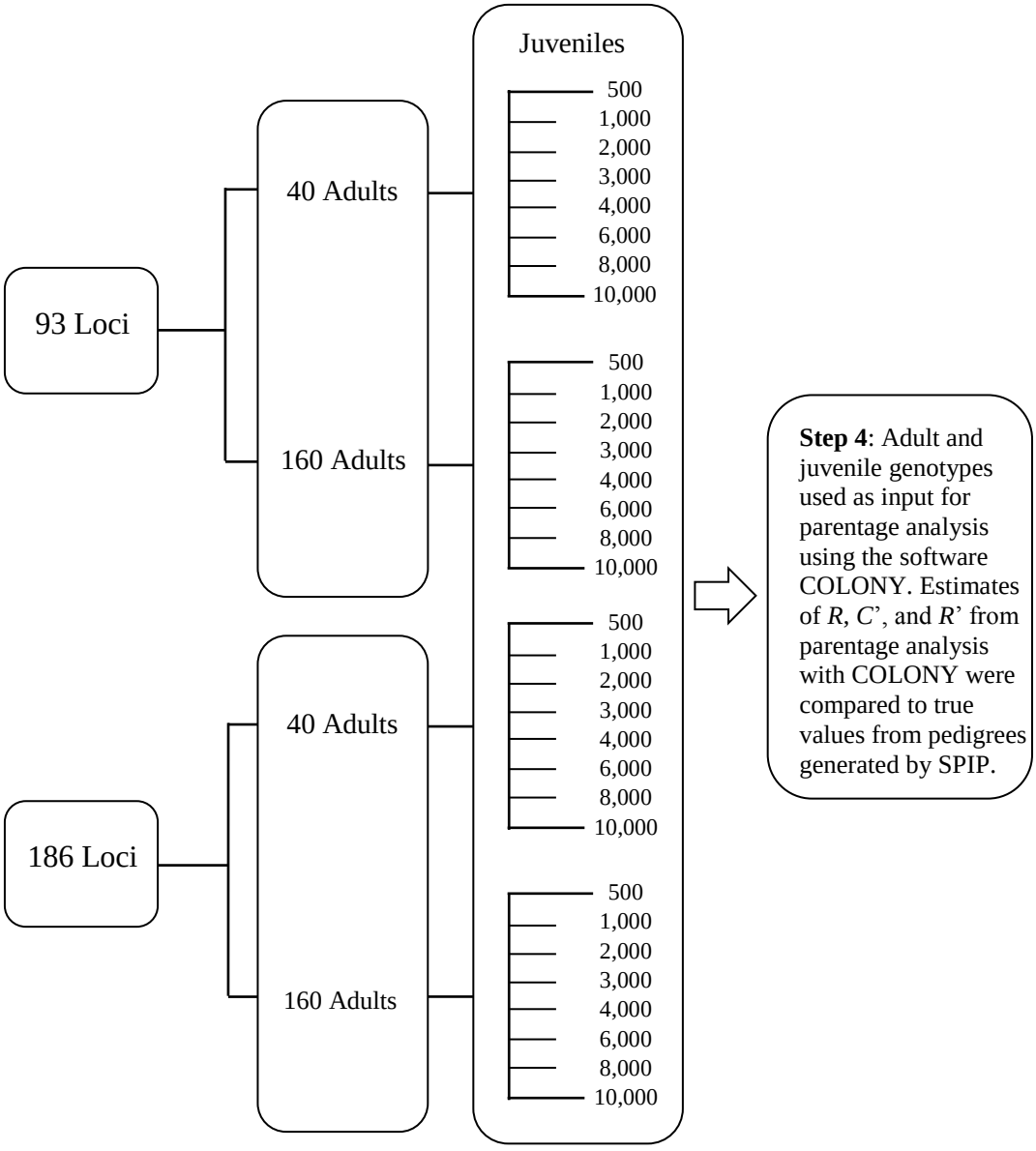
traditional escapement estimation methods (e.g., area-under-the-curve, Jolly-Seber, and redd counts) and their precision evaluated in comparison to standards set by McElhany et al. (2000), no evaluation has been conducted on the accuracy of genetic parentage analysis used to estimate *R*, *C*', and *R*', which have been assumed to be measured without error in applications of tGMR. To address this issue, I simulated an age-structured population of coho salmon (*Oncorhynchus kisutch*), where multiple generations participated in spawning. Herein, males two years of age will be considered adults and included in the analysis. Simulations also generated individual multilocus genotypes and pedigrees for all individuals were recorded. Multilocus genotypes were simulated for each individual based on the allele frequencies of a coho salmon population in northern California. I then emulated the field application of tGMR by taking a simple random sample (without replacement) of *M* simulated adults to represent field collection of carcasses during the first sampling occasion in tGMR. Next, a simple random sample (without replacement) of *n* juveniles resulting from these adults was selected to represent the second sampling occasion in tGMR. The multilocus genotypes of the random samples of adults and juveniles were then subjected to parentage analyses using the software COLONY. This allowed comparison of estimates of *R*, *C'*, and *R'* from parentage analysis to the true values taken from pedigrees for the simulated population. I evaluated the effect of adult sample size (*M*= 40 and 160), juvenile sample size (*n*= 500, 1000, 2000, 3000, 4000, 6000, 8000, and 10000), and locus number (93 and 186) on accuracy of estimating *R*, *C'*, and *R'*.

METHODS

To evaluate the accuracy of genetic pedigree reconstructions used for tGMR I conducted an analysis with the following steps (Figure 1): (1) Simulated pedigrees and multilocus genotypes (for 93 and 186 loci, respectively) were generated for two age-structured populations (multiple generations participate in spawning) of coho salmon (adult males of age 2, 3, and 4; adult females of age 3 and 4) using the software SPIP (Simulate Pedigrees In a Population; Anderson and Dunham 2005), (2) for each of these two settings, a simple random sample (without replacement) of adults from the simulated population was selected to emulate the first sampling session in tGMR, (3) for each of these two settings, a simple random sample (without replacement) of juveniles was selected to represent the second sampling occasion in tGMR, (4) the software COLONY (Jones and Wang 2010) was used to reconstruct pedigrees based on the multilocus genotypes from the sampled parents and offspring. Accuracy was assessed by comparing estimated parameters required for the "with replacement" method (R) and "without replacement" method (C' and R') to the true values taken from known simulated pedigrees from the two populations. Simulations were run for 32 combinations of loci (93 SNPs and 186 SNPs), parent subsample size (M= 40 and 160), and juvenile sample sizes (n= 500, 1000, 2000, 3000, 4000, 6000, 8000, and 10000). For each combination of loci, and sample sizes for parents and juveniles, 100 replicate runs of steps three and four were completed for a total of 3200 total runs.

Population Simulation

The software SPIP was used to simulate an age-structured population of coho salmon (Figure 1; Step 1). This population was modeled based upon life history and demography of coho salmon in Freshwater Creek, Humboldt County, California. The adult population of 800 individuals had a fixed cohort size of 14,500 juveniles, an equal adult sex ratio, and the majority of spawners were of age 3. The number of offspring contributed to a cohort by an individual parent followed a skewed distribution. To reach the target of 800 adult spawners, SPIP was run repeatedly until the target adult population size was reached. The software SPIP was also used to generate multilocus genotypes and pedigrees for individuals in the simulated population. The population was simulated for 40 years; parent genotypes were taken from year 39 and juveniles from year 40. The input parameters for SPIP are in the Appendix. Allele frequencies for 93 polymorphic SNP loci from the 2012-2013 returning adults in Freshwater Creek were repeated to create an input file with 186 loci for simulating adult and juvenile multilocus genotypes (Supplementary material, Table 1). A simulated population was created using the 186 loci input file. The genotypic data from the first 93 loci was harvested from the simulated population, allowing the creation of two populations (93 and 186 loci), but with identical demography. There was no missing data in the multilocus genotypes in the simulated individuals.



Step 1: Simulation of two populations with either 93 or 186 SNP loci using the software SPIP.

Step 2: Simple random sampling of simulated adult genotypes to resemble the first sampling occasion in tGMR.

Step 3: Simple random sampling of offspring genotypes to mimic the second sampling occasion in tGMR. A total of 100 replicate samples were selected while keeping the number of loci and adults fixed.

Figure 1. Diagrammatic representation of the simulation procedures used to evaluate accuracy of parentage analysis used in transgenerational mark-recapture.

Adult and Juvenile Subsampling

To emulate field collection of adults (e.g., via carcass surveys) during the first sampling occasion in tGMR, a simple random sample of either 40 or 160 adults were selected from the simulated adult population of 800 (Figure 1; Step 2). Adult sample numbers were chosen to reflect a realistic sampling level for a small coho salmon population, where adult carcasses are scarce and recovery of carcasses during field surveys is unlikely to exceed 20% of total escapement (Seth Ricker, California Department of Fish and Wildlife, pers. comm.). To represent the second sampling occasion, a simple random sample of juveniles was selected from the offspring population of 14,500. Juvenile samples sizes were of 500, 1000, 2000, 3000, 4000, 6000, 8000, and 10000 individuals. Sample sizes were selected based upon field observations that suggest collections of large numbers of juveniles are theoretically feasible.

Parentage Analysis

To reconstruct pedigrees, the multilocus genotypes of the parents and juvenile samples from the simulated population were used as input to the maximum likelihood algorithms (Wang 2004; Wang and Santure 2009; Wang 2012; Wang 2013) implemented in COLONY (version 2.0.6.1) (Jones and Wang 2010), to produce estimates of R, C', and R' used for tGMR. COLONY uses multilocus genotypes to infer sibship and parentage among individuals using a full-pedigree likelihood method. COLONY has the ability to reconstruct full- and half-sibling family clusters from juvenile genotypes, allowing

inference concerning “unmarked” parents that were not in the original sample of marked (genotyped) candidate fathers or mothers. The settings for parentage analysis using COLONY were: a polygamous mating system, no inbreeding or clones, species are dioecious and diploid, length of run and likelihood precision were set to medium, full likelihood analysis, no sibship priors, genotyping error rates are 0.0001 per marker, and unknown allele frequency. COLONY’s computation time is substantial. It was estimated that it would have taken 5643 days of computation time to complete the 3200 simulation runs using a single computer (assuming an Intel i7 dual core processor with 16 GB RAM). To address this limitation, simulations were ran in parallel using between 400 and 1200 cores on the FARM II cluster at the University of California, Davis campus.

Accuracy Assessment

The initial coho salmon population simulation using the software SPIP included pedigrees and allowed determination of the true values of *R*, *C*’ and *R*’ for each simulated juvenile sample. These values were compared to estimated values generated based upon parentage analysis with COLONY (Figure 1; Step 4), because the magnitude of the expected values of *R*, *C’*, and *R’* vary with scenario (loci number, adult sample size and random sample of selected juveniles), assessment of the performance of COLONY was more complex than for a typical situation.

The following calculations were used to assess the performance of COLONY for estimation of R. Let $R(s) =$ SPIP- simulated known total number of times adults in the first sampling occasion (M) are assigned as parents of individuals from the juvenile

sample s, $s = 1, 2, 3, \ldots, 100$. Let $\bar{R} = \sum_{s=1}^{100} \frac{R(s)}{100}$ = average SPIP-simulated known total number of times adults in the first sampling occasion (*M*) are assigned as parents of individuals from the juvenile sample s. Let $\hat{R}(s)$ = COLONY's estimate of $R(s)$ for a given SPIP output. Let $\bar{\hat{R}} = \sum_{s=1}^{100} \frac{\hat{R}(s)}{100}$ = COLONY's average estimate of $R(s)$. A measure of average bias was calculated as Bias($\hat{R}$) $= \bar{\hat{R}} - \bar{R}$ and proportional bias was calculated as:

$$\text{PropBias}(\hat{R}) = \text{Bias}(\hat{R})/\bar{R}$$

Approximate sampling variance of $\hat{R}$ was calculated as:

$$V(\hat{R}) = \sum_{s=1}^{100} (\hat{R}(s) - \bar{\hat{R}})^2/(100 - 1)$$

and mean squared error using:

$$MSE(\hat{R}) = \sum_{s=1}^{100} (\hat{R}(s) - \bar{R})^2/(100 - 1)$$

The following calculations were used to assess COLONY's performance for estimation of C'. Let $C'(s)$ = SPIP- simulated known number of parents who gave rise to the juvenile subsample s, $s = 1, 2, 3, \ldots, 100$. Let $\overline{C'} = \sum_{s=1}^{100} \frac{C'(s)}{100}$ = average SPIP-simulated known number of parents who gave rise to the juvenile subsample s. Let $\widehat{C'}$ = COLONY's estimate of $C'(s)$ for a given SPIP output (juvenile sample). Let $\overline{\widehat{C'}}$ = $\sum_{s=1}^{100} \frac{\widehat{C'}(s)}{100}$ = COLONY's average estimate of $C'(s)$ over 100 independent juvenile samples. Then, Bias($\widehat{C'}$) $= \overline{\widehat{C'}} - \overline{C'}$ and proportional bias was calculated as:

$$\text{PropBias}\left(\widehat{C'}\right) = \text{Bias}(\widehat{C'})/\overline{C'}$$

Approximate sampling variance of $\widehat{C'}$ was calculated as:

$$V\left(\widehat{C'}\right) = \sum_{s=1}^{100}(\widehat{C'}(s) - \overline{\widehat{C'}})^2/(100-1)$$

and mean squared error as:

$$MSE\left(\widehat{C'}\right) = \sum_{s=1}^{100}(\widehat{C'}(s) - \overline{C'})^2/(100-1)$$

Completely analogous equations were used to characterize bias, proportional bias, sampling variance, and mean square error of *R*'.

RESULTS

Population Simulation

In the simulated coho salmon population generated with the software SPIP a total of 745 of the 800 adults produced at least one offspring (93% reproductive success). The ratio of reproductively successful males to females was about 1:1 (Males= 381, Females= 364). The number of juveniles per male ranged from 1 to 193 (mean= 38.36, variance = 1,143.89), following a skewed distribution (Figure 2A). The number of juveniles per female ranged from 13 to 78 (mean= 39.83, variance= 160.77), following a symmetrical distribution (Figure 2B). The number of juveniles per reproductively successful adult (*N*= 745, males and females combined) ranged from 1 to 193 (mean= 39.89, variance= 689.53), following a skewed distribution (Figure 2C). The distribution of juvenile production for males and females within the adult subsamples (*M*) was similar to that of the full simulated population (Table 1).

Table 1. The percentage of adults that produced at least one offspring (reproductive success (RS)), and the number of offspring per male and female (range, mean, and sampling variance) for two subsamples of adults (*M*= 40, 160) and the entire simulated population (*N*= 800).

		Offspring/ Male			**Offspring/ Female**		
Adults	**RS%**	**Range**	**Mean**	**Variance**	**Range**	**Mean**	**Variance**
40	93	1 – 142	40.58	1982.37	23 – 54	38.67	86.71
160	91	1 – 139	37.42	1159.70	15 – 78	40.56	195.65
800	93	1 – 193	38.36	1143.89	13 – 78	39.83	160.77

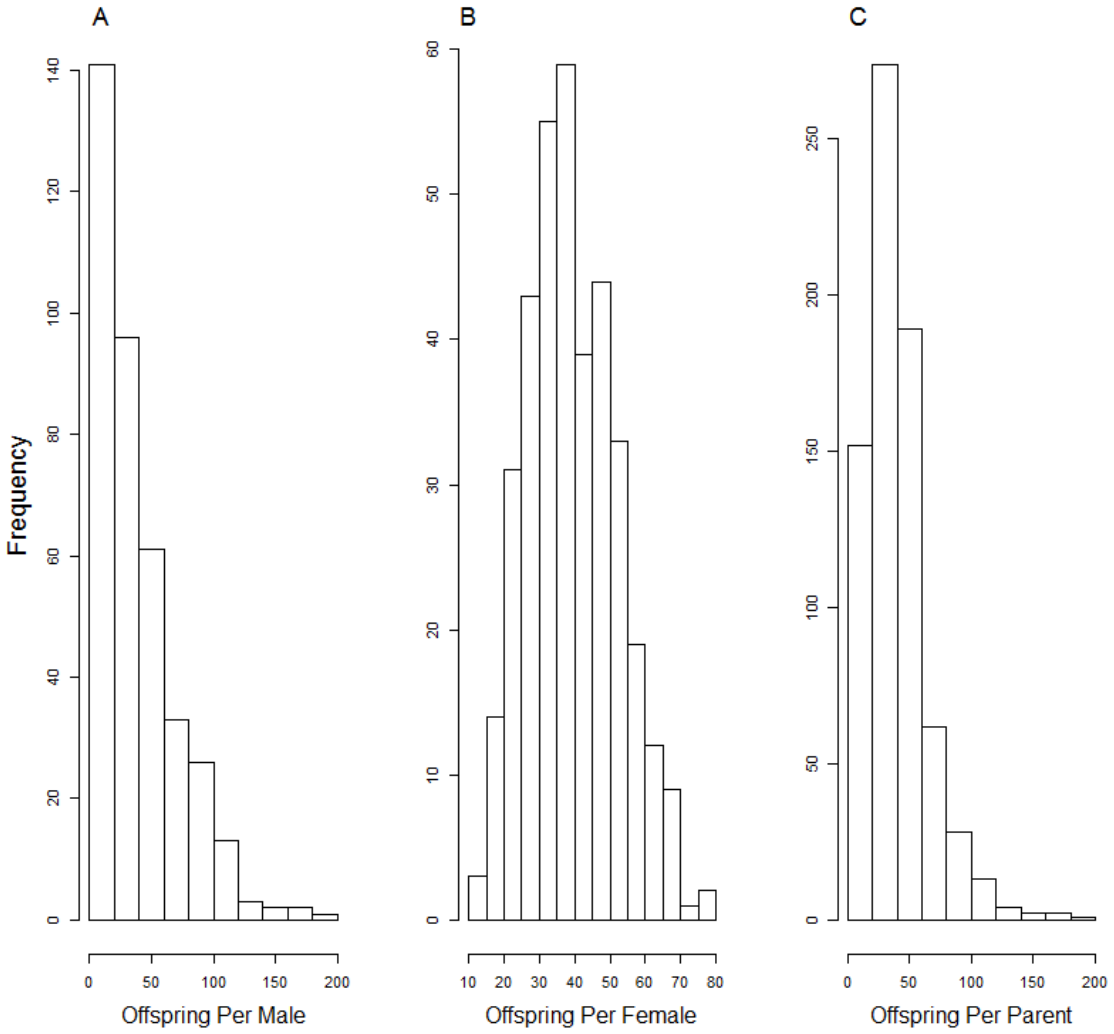


Figure 2. The number of offspring per adult male parent (A), per female parent (B), and both male and female parents combined (C). Distributions are for the simulated universe generated using the software SPIP, in which 745 of the 800 adults produced at least one offspring.

Accuracy Assessment

For the "with replacement" model, parentage analysis was used to estimate the total number of times adults from the first sampling occasion (*M*) were assigned as parents of individuals from the juvenile sample, or the number of recaptures (*R*). Proportional bias of estimates of *R* ranged from 0.05 to -0.06 depending upon loci number, parent sample, and juvenile sample size (Table 2; Figure 3). At smaller offspring sample sizes (*n*= < 2,000), using 186 loci provided approximately unbiased estimates of *R* whereas using 93 loci resulted in biased estimates of *R* (ranging from 0.00 to -0.05). In contrast, at larger offspring sample sizes (*n*= > 2,000) estimates of *R* were always negatively biased (range -0.01 to -0.06) regardless of the number of loci, parents, or offspring. However, the analysis did indicate that increased adult sample size and loci number may reduce biased estimates of *R* at larger offspring sample sizes (*n*= > 2000) (Figure 3).

For the "without replacement" model, proportional bias of $\widehat{C'}$ ranged from -0.40 to 0.23 (Table 3). Juvenile sample size of 2000, using 186 loci, generated the lowest proportional bias (0.01 with 160 parents, and -0.02 with 40 parents). When using 93 loci, proportional bias was lowest at a juvenile sample size of 3000 (0.01 for 40 parents and 0.02 for 160 parents). Regardless of the number of loci used and parent sample size, proportional bias was negative at relatively small juvenile sample size ($n \leq 2{,}000$) and became positive once juvenile sample sizes became large ($n \geq 3{,}000$). Increasing parent sample size and the number of loci decreased bias when juvenile sample size was less

than about 3000 individuals, but bias in *C*' did not decline to zero with increasing juvenile sample size.

COLONY's estimates of *R'* exhibited small bias when using 93 loci, 40 parents, and juvenile sample sizes of ≤ 2,000 (0.01 to 0.03) but in the remainder of scenarios estimates of *R*' were approximately unbiased (Table 4; Figure 5).

Table 2. Simulation results for $\hat{R}$, the estimated total number of times adults in the first sampling occasion (*M*) are assigned by COLONY as parents of individuals from the juvenile sample, or the number of recaptures. $\bar{R}$ is the mean true value taken from pedigrees of the simulated coho salmon population and $\bar{\hat{R}}$ represents the mean number of recaptures estimated by parentage analysis with COLONY. Reported are the proportional bias (PropBIAS($\hat{R}$)), sampling variance (VAR($\hat{R}$)), and mean square error(MSE($\hat{R}$)) for $\hat{R}$.

Loci	*Parents*	*Juveniles*	$\bar{R}$	$\bar{\hat{R}}$	*PropBIAS*($\hat{R}$)	*VAR*($\hat{R}$)	*MSE*($\hat{R}$)
93	40	500	49.78	52.49	0.05	42.68	50.09
		1000	101.34	104.18	0.03	90.23	98.38
		2000	202.19	203.00	0.00	207.31	207.98
		3000	304.68	302.51	-0.01	200.80	205.55
		4000	401.67	395.01	-0.02	287.38	332.19
		6000	607.38	587.42	-0.03	533.64	936.07
		8000	807.69	770.74	-0.05	775.39	2154.48
		10000	1009.17	961.54	-0.05	1147.42	3438.95
93	160	500	194.57	199.56	0.03	193.48	218.63
		1000	389.83	395.46	0.01	331.42	363.44
		2000	781.13	780.55	0.00	587.40	587.74
		3000	1177.66	1170.02	-0.01	586.65	645.61
		4000	1559.19	1538.18	-0.01	987.97	1433.85
		6000	2335.97	2286.71	-0.02	1645.50	4096.56
		8000	3115.79	3030.67	-0.03	1634.41	8953.01
		10000	3897.01	3769.12	-0.03	2835.60	19356.66
186	40	500	50.17	50.41	0.00	43.25	43.31
		1000	100.81	101.02	0.00	71.15	71.20
		2000	201.59	200.68	0.00	146.50	147.34
		3000	302.92	298.54	-0.01	189.79	209.16
		4000	404.37	394.17	-0.03	307.80	412.89
		6000	607.22	578.86	-0.05	695.72	1508.13
		8000	806.69	761.29	-0.06	1232.57	3314.55
		10000	1013.86	952.88	-0.06	1329.60	5085.72
186	160	500	194.53	194.94	0.00	125.01	125.18
		1000	387.94	387.99	0.00	316.15	316.15
		2000	779.71	777.43	0.00	589.30	594.55
		3000	1160.64	1148.68	-0.01	775.17	919.66
		4000	1560.40	1533.47	-0.02	813.00	1545.55
		6000	2343.33	2280.58	-0.03	1524.55	5501.88
		8000	3118.18	3003.41	-0.04	2411.50	15716.70
		10000	3892.84	3721.88	-0.04	3888.33	15716.70

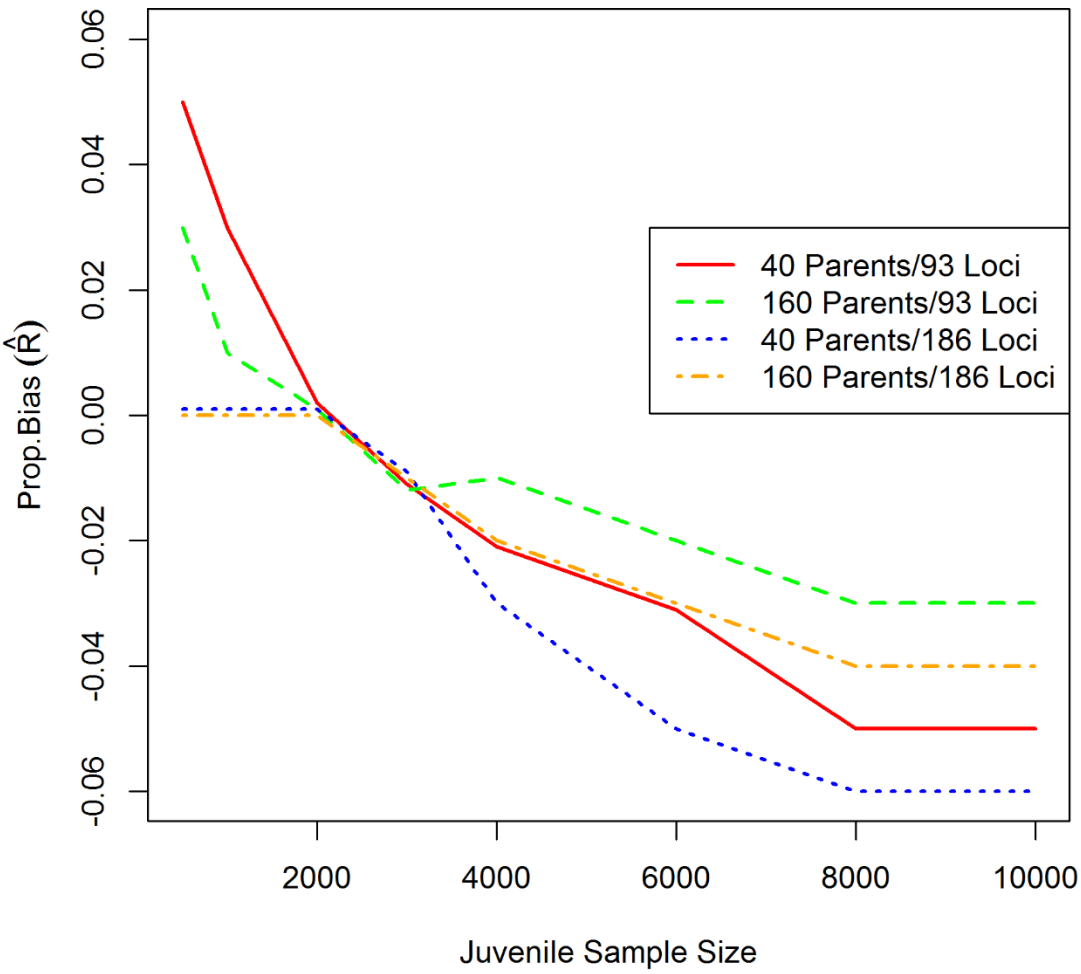


Figure 3. Proportional bias of $\hat{R}$, the estimated total number of times adults in the first sampling occasion (M) are assigned as parents of individuals from the juvenile sample, or the number of recapture.

Table 3. Simulation results for $\widehat{C'}$, the estimated total number of distinct parents that gave rise to the juvenile subsample based on COLONY software. $\overline{C'}$ is the true value taken from simulated pedigrees and $\overline{\widehat{C'}}$ represents the average total number of unique parents that gave rise to the juvenile subsample estimated by parentage analysis with COLONY. Reported are the proportional bias (PropBIAS($\widehat{C'}$)), sampling variance (VAR($\widehat{C'}$)), and mean squared error (MSE($\widehat{C'}$)) for C'.

Loci	*Parents*	*Juveniles*	$\overline{C'}$	$\overline{\widehat{C'}}$	*PropBIAS*($\widehat{C'}$)	*VAR*($\widehat{C'}$)	*MSE*($\widehat{C'}$)
93	40	500	487.42	290.02	-0.40	41.51	39401.88
		1000	616.30	457.57	-0.26	67.80	25517.51
		2000	682.74	629.85	-0.08	70.39	2896.00
		3000	702.68	712.06	0.01	99.01	187.88
		4000	713.78	761.11	0.07	84.00	2346.75
		6000	725.12	825.92	0.14	206.88	10470.16
		8000	732.50	863.06	0.18	239.92	17458.01
		10000	737.21	880.78	0.19	260.76	21081.31
93	160	500	488.65	324.57	-0.34	55.38	27249.57
		1000	615.47	491.02	-0.20	76.69	15720.93
		2000	683.05	648.74	-0.05	75.63	1264.70
		3000	702.24	718.28	0.02	94.85	354.73
		4000	713.41	767.88	0.08	97.38	3094.33
		6000	725.36	820.62	0.13	209.67	9375.80
		8000	732.09	849.13	0.16	192.17	14028.90
		10000	737.44	877.01	0.19	184.07	19860.62
186	40	500	488.83	346.31	-0.29	46.38	20563.50
		1000	614.16	515.23	-0.16	78.60	9964.61
		2000	683.48	667.22	-0.02	77.85	344.91
		3000	702.90	735.00	0.05	84.65	1125.46
		4000	713.70	781.98	0.10	100.81	4810.06
		6000	725.17	840.21	0.16	145.88	13513.77
		8000	732.04	876.45	0.20	409.93	21474.82
		10000	737.42	906.76	0.23	223.78	29189.47
186	160	500	489.24	381.17	-0.22	48.47	11845.56
		1000	614.45	544.17	-0.11	101.98	5091.15
		2000	682.41	678.63	-0.01	58.52	72.95
		3000	703.28	740.14	0.05	86.34	1458.73
		4000	713.66	780.55	0.09	102.25	4621.72
		6000	725.40	828.86	0.14	216.97	11029.06
		8000	732.32	859.17	0.17	236.49	16489.94
		10000	737.03	896.68	0.22	272.44	16489.94

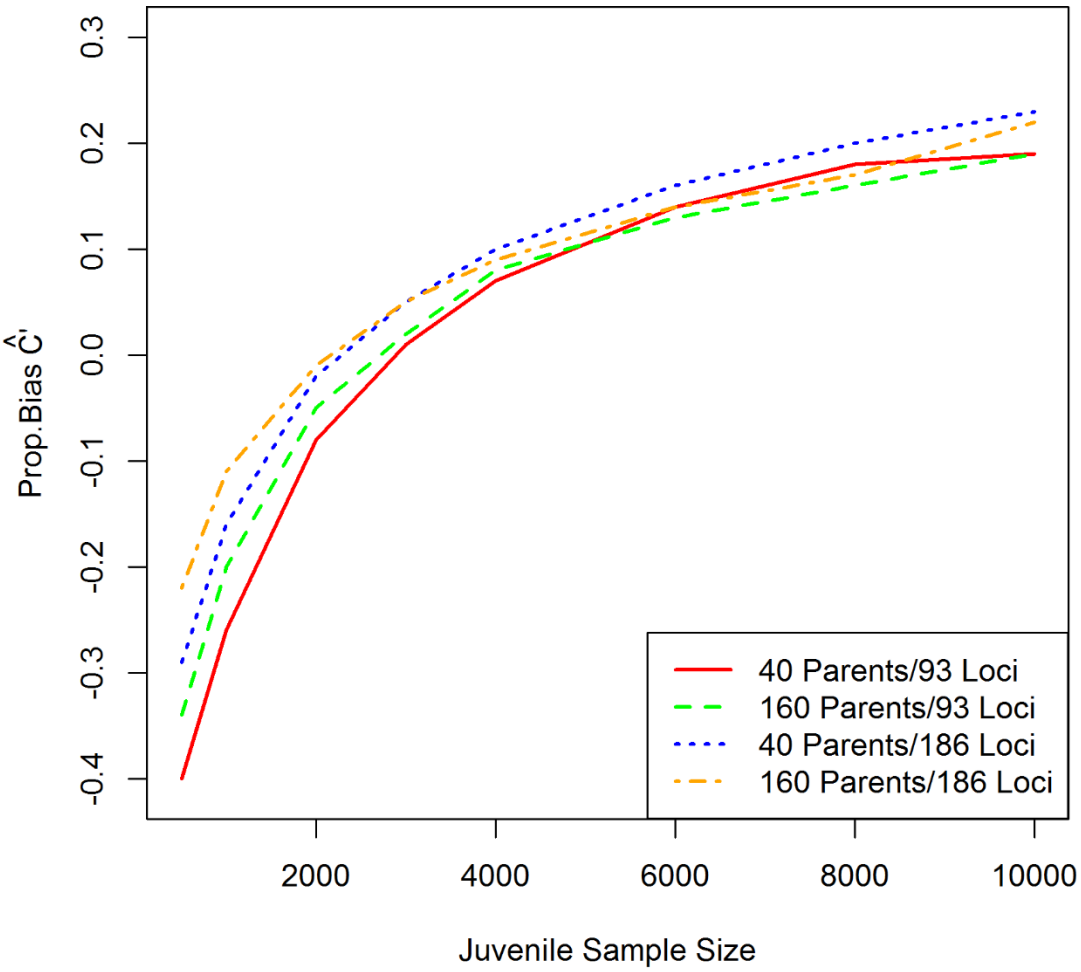


Figure 4. Proportional bias of $\widehat{C'}$, the estimated total number of distinct parents that gave rise to the juvenile subsample.

Table 4. Simulation results for $\widehat{R'}$, the estimated total number of distinct adults from the first sampling occasion (*M*) that were responsible for producing at least one offspring in the juvenile subsample(number of recaptures). $\overline{R'}$ is the true value taken from simulated pedigrees and $\overline{\widehat{R'}}$ represents the average number of recaptures estimated by parentage analysis with COLONY. Reported are the proportional bias (PropBIAS($\widehat{R'}$)), sampling variance (VAR($\widehat{R'}$)), and mean square error(MSE($\widehat{R'}$)) for R'.

Loci	*Parents*	*Juveniles*	$\overline{R'}$	$\overline{\widehat{R'}}$	*PropBIAS*($\widehat{R'}$)	*VAR*($\widehat{R'}$)	*MSE*($\widehat{R'}$)
93	40	500	23.53	24.16	0.03	5.47	5.87
		1000	29.71	30.14	0.01	2.53	2.71
		2000	32.70	32.93	0.01	1.70	1.76
		3000	33.65	33.74	0.00	1.00	1.01
		4000	34.20	34.28	0.00	0.99	1.00
		6000	34.94	34.98	0.00	0.95	0.95
		8000	35.26	35.28	0.00	0.81	0.81
		10000	35.87	35.90	0.00	0.60	0.60
93	160	500	94.64	94.23	0.00	27.35	27.52
		1000	118.60	119.16	0.00	13.49	13.81
		2000	130.77	131.10	0.00	5.38	5.49
		3000	134.28	134.53	0.00	5.08	5.14
		4000	136.89	137.07	0.00	3.50	3.53
		6000	139.10	139.15	0.00	3.00	3.00
		8000	141.10	141.13	0.00	2.86	2.86
		10000	142.59	142.70	0.00	1.99	2.00
186	40	500	23.39	23.48	0.00	4.84	4.85
		1000	29.49	29.51	0.00	3.06	3.06
		2000	32.83	32.83	0.00	1.17	1.17
		3000	33.68	33.68	0.00	1.19	1.19
		4000	34.20	34.20	0.00	0.87	0.87
		6000	34.91	34.91	0.00	1.05	1.05
		8000	35.34	35.34	0.00	0.81	0.81
		10000	36.04	36.04	0.00	0.56	0.56
186	160	500	94.43	94.59	0.00	23.82	23.85
		1000	118.19	118.23	0.00	11.41	11.41
		2000	130.71	130.70	0.00	5.38	5.38
		3000	134.81	134.81	0.00	4.50	4.50
		4000	136.76	136.81	0.00	3.43	3.43
		6000	139.16	139.16	0.00	3.25	3.25
		8000	141.27	141.27	0.00	2.64	2.64
		10000	142.51	142.51	0.00	1.97	2.64

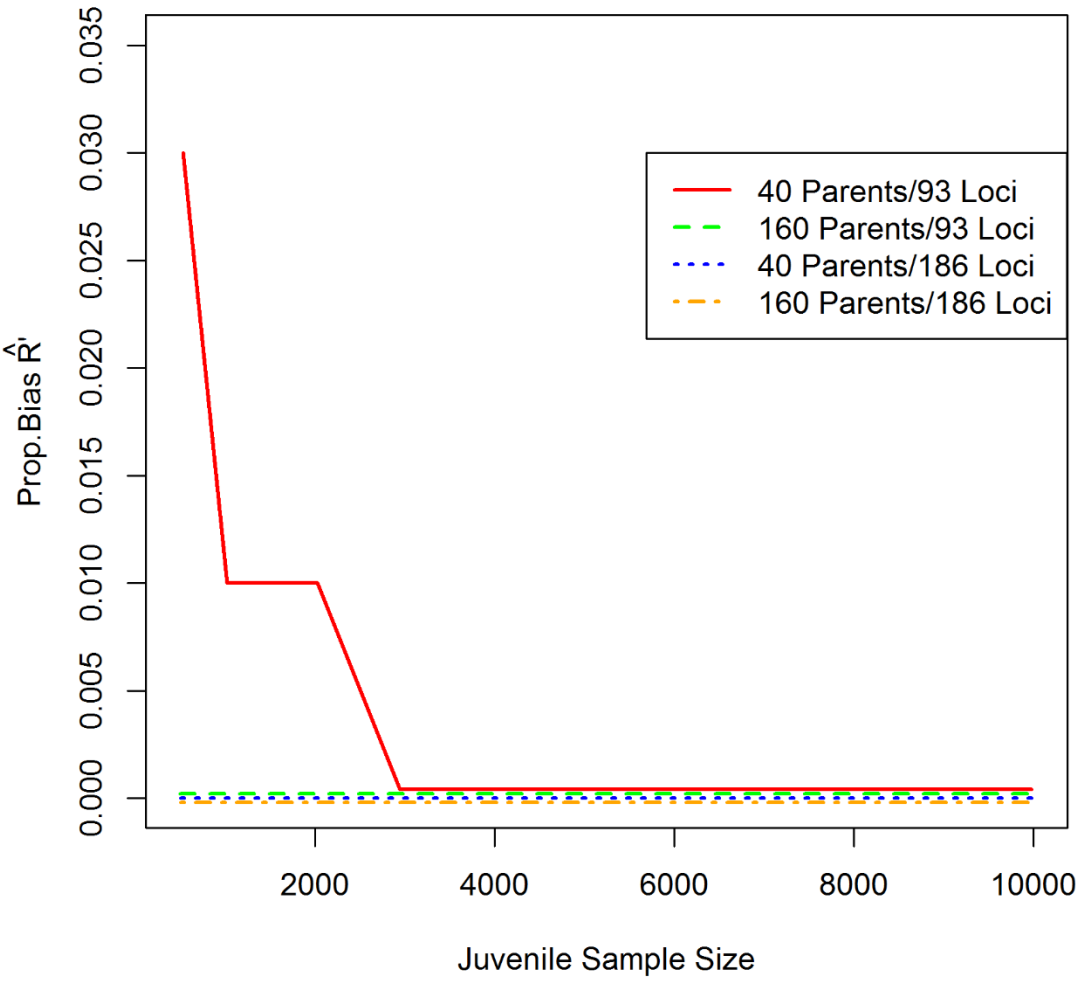


Figure 5. Proportional bias of $\widehat{R'}$, total number of distinct adults from the first sampling occasion (M) that were responsible for producing at least one offspring in the juvenile subsample

DISCUSSION

Accurate estimation of the tGMR variables (*R*, *C*', and *R*') is needed for use of tGMR estimators, in addition to the requirement to meet tGMR assumptions (see Chapter 2, *Discussion*). Overall, findings of this study indicated that COLONY's estimates of *R* had modest proportional bias ranging from 0.05 to -0.06. COLONY's estimates of *C*' had more serious proportional bias ranging from -0.40 to 0.23. COLONY's estimates of *R*' had negligible proportional bias (0 to 0.01) except when the number of loci was small (*k*= 93), adult sample size was small (*M*= 40), and juvenile sample size was small (*n*= 500) in which case proportional bias was 0.03.

Previous work has shown that in some cases parentage analysis with COLONY may create larger families than actually exist, referred to as "clumping" or Type 1 errors (Wang 2013) (Figure 6A). When parentage analysis suffers from "clumping" it would introduce positive bias in the number of recaptures (*R*). Evidence of "clumping" was present with 93 loci and relatively small juvenile sample sizes (*n*= ≤ 2,000) but absent with 186 loci and relatively small juvenile sample sizes (*n*= ≤ 2,000). This implies that 186 loci was enough information for COLONY to correctly identify juvenile-parent pairs when those parents were present in the adult sample (*M*) and when the juvenile sample size was ≤ 14% of the total population (*PropBIAS* = 0). In contrast, once the juvenile sample size exceeded 2,000 individuals there was evidence for "splitting" (Type 2 errors; Wang 2013) or creation of many small families by parentage analysis (Figure 6B). "Splitting" introduced negative bias to estimates of *R* and was observed under all

combinations of loci (*k*= 93, 186), adults (*M*= 40, 160) when juvenile sample size was >2000. When families were split (Type 2 error), fewer parent-offspring relationships were inferred among the sampled individuals. This negative bias was reduced by increasing the adult sample size (*M*= 160). Increasing the adult sample size had more influence than the juvenile sample size on reducing the occurrence of both types of errors. The larger number of adult genotypes available allows COLONY to detect more adults present in *M* through juvenile genotypes while inferring full- and half-sibling relationships. Ultimately, inaccuracy in estimates of *R* affects the total escapement. On average, negative bias in estimation of *R* will inflate estimates of *N*, while positive bias in estimation of *R* will have the opposite effect on estimates of *N*.

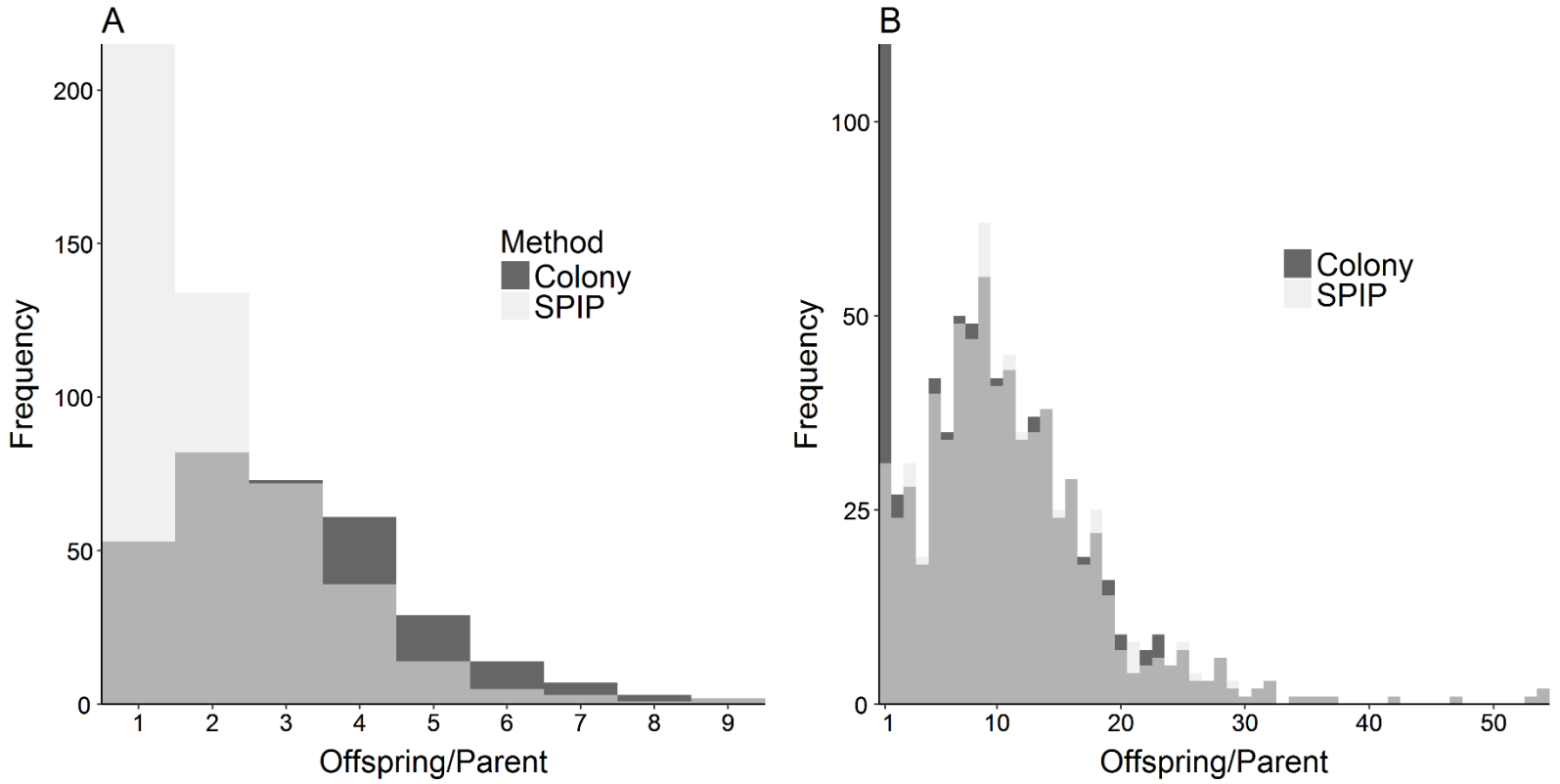


Figure 6. The frequency distribution of the number of offspring per parent at 93 SNP loci with an adult sample size of 40 individuals. A subsample of 500 juveniles (A) is compared to a subsample of 4,000 juveniles (B). SPIP is the known frequency distribution of the number of offspring per parent and Colony is the estimated frequency distribution of the number of offspring per parent

COLONY's estimates of *C*' varied dramatically and had substantial proportional bias except for a very limited set of conditions (*k*= 93, *n*= 3000 and *k*= 186, *n*= 2000). The parameter *C*' consists of two parts: the number of distinct "marked" (genotyped) adults (*R*') from *M* that are assigned as parents of juveniles and the number of additional distinct unmarked parents (*C*'- *R*') inferred to be represented among juveniles based on parentage analysis. Our analysis indicates that parentage analysis generally provides accurate estimates of *R*', which were essentially unbiased unless sample sizes were small (*k*= 93, *M*= 40, and *n*= < 2000). Thus, errors in estimation of *C'* must result primarily from an inability to accurately infer the additional (unmarked) parents (*C*'- *R*') required to describe the offspring sample. Our findings suggest that differences in juvenile sample size may result in a large degree of variability in estimating *C'*. When juvenile sample sizes fell below 3,000 individuals, Type 1 errors were introduced, creating a negative bias for *C'*. However, when juvenile sample sizes exceeded 3,000 individuals, Type 2 errors were observed, which created a positive bias for *C'*. The variability in sign and magnitude of proportional bias associated with COLONY's estimation of *C'* raises serious concerns for use of the "without replacement" model. Type 1 errors would typically result in negative bias in an estimate of total escapement, whereas Type 2 errors would typically result in positive bias in estimating total escapement. Table 3 and Figure 4, do clearly show a point (*n*= 2000, 3000) where there is no apparent substantial proportional bias in *C'* regardless of the number of loci used or adult sample size. However, it is impossible to know, without additional simulations, if this result would be obtained for a different adult population size and a different total smolt production.

In contrast to difficulties associated with estimation of *C*', parentage analysis using COLONY was accurate when assigning juveniles back to their putative parents (*R'*). There were low levels of positive proportional bias (0.01 – 0.03) but only when the least amount of genetic information was available (*k*= 93, *M*= 40, and *n*= < 2000). As the sample sizes increased above these levels, no bias was evident in estimates of *R*' obtained using parentage analysis.

In general, it appears that the performance (proportional bias) of COLONY for estimation of *R* and *C*' depends sensitively on size of the juvenile sample. COLONY's accuracy apparently deteriorates if parents produce many families of size one ("singletons") and/or when there are large families (many juvenile per parent pair) produced by spawning adults. For example, there is a higher frequency of small families than large families within the simulated SPIP universe, as is expected for salmonid populations (Figure 2C). When the juvenile sample size was 500 out of 14,500 individuals, it was more likely that many unrelated individuals "singletons" were selected. Many of these unrelated individuals were probably then clumped, forming fewer families than actually existed. In contrast, when the juvenile sample size was very large, say 10,000 out of 14,500 individuals, it was more likely that related individuals were chosen resulting in fewer "singletons". These "larger" families are likely then split, forming fewer families of large size than should exist. These errors are expected with varying family sizes as clumping occurs from many small families being present or splitting when many large families are present (Wang 2013).

Type 1 and 2 errors are evident in all of COLONY's estimates of tGMR parameters. The degree of "clumping" or "splitting" is dependent on the number of loci used and sample sizes of both adults and juveniles. As expected based on previous research, COLONY falsely formed families when juvenile samples were small (Wang 2013). However, COLONY falsely split large families of related individuals when sample sizes were relatively high (Almudevor and Anderson 2012). At 93 loci, no missing data, a mean MAF of 25% (q = 0.25), and a per locus genotyping error rate of 0.0001%, I was able to accurately detect our sampled adult genotypes through the subsample of juveniles. However, even when I doubled the number of loci to 186, I was unable to eliminate Type 1 and 2 errors when inferring full- and half-sibling relationships. Increasing the number of loci, adult sample size, and juvenile sample size did not smoothly reduce absolute proportional bias associated with estimating *R* and *C'* but instead changed the sign of bias from negative to positive. Reliable identification of parent-offspring pairs may require hundreds of polymorphic SNP loci and thousands may be needed to correctly infer half-siblings (Bravington et al. 2016).

I conclude that even with no missing data, 93 and 186 loci are not powerful enough to accurately estimate *R*, *C*' and *R*' for implementation of tGMR. Further investigation is needed on whether or not the point where no bias is introduced for *R* and minimal bias is introduced to *C'* is relevant to the juvenile sample size itself or the proportion of the population the sample represents. To address this, multiple populations would need to be simulated at different sizes and simulations conducted at juvenile sample sizes of 2,500 individuals or ~14% of the juvenile population regardless of total

juvenile population size. This would allow us to determine if a fixed sample (e.g., 2,000 individuals) at 93 or 186 loci allows COLONY to accurately infer full- and half-siblings or if the proper juvenile sample size should be a given proportion of the population. Simulations could also be run with a much greater number of SNP loci, for example at a suggested number of hundreds to thousands of SNP loci. This may give us insight on the minimum number of SNP loci needed to eliminate Type 1 and 2 errors. The availability of multiallelic loci where each loci has the informative power of microsatellite markers, may provide improved power for parentage analysis.

LITERATURE CITED

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

Almudevar, A., and E. C., Anderson. 2012. A new version of PRT software for sibling groups reconstruction with comments regarding several issues in the sibling reconstruction problem. Molecular Ecology Resources 12(1):164-178.

Anderson, E. C. 2010. Computational algorithms and user-friendly software for parentage-based tagging of Pacific salmonids. Final report to the Pacific Salmon Commission's Chinook Technical Committee (US Section). 46 pp.

Anderson, E. C., and K. K. Dunham. 2005. spip 1.0: a program for simulating pedigrees and genetic data in age-structured populations. Molecular Ecology Notes 5(2):459-461.

Anderson, E.C., and J. C. Garza. 2006. The power of single-nucleotide polymorphisms for large-scale parentage inference. Genetics 172(4):2567-82.

Blouin, M. S. 2003. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. Trends in Ecology & Evolution 18(10):503-511.

Bravington, M. V., H. J. Skaug, and E. C. Anderson. 2016. Close-Kin mark recapture. Statistical Science 31(2):259-274.

Fiumera, A. C., B. A. Porter, G. D. Grossman, and J. C. Avise. 2002. Intensive genetic assessment of the mating system and reproductive success in a semi-closed population of the mottled sculpin, *Cottus bairdi*. Molecular Ecology 11(11):2367-2377.

Israel, J. A., and B. May. 2010. Indirect genetic estimates of breeding population size in the polyploid green sturgeon (*Acipenser medirostris*). Molecular Ecology 19(5):1058-1070.

Jones, A. G., and J. C. Avise. 1997. Microsatellite analysis of maternity and the mating system in the Gulf pipefish *Syngnathus scovelli*, a species with male pregnancy and sex-role reversal. Molecular Ecology 6(3):203-213.

Jones, O. R., and J. Wang. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources 10(3):551-555.

McElhany, P., M. H. Ruckelshaus, M. J. Ford, T. C. Wainwright, and E. P. Bjorkstedt. 2000. Viable salmonid populations and the recovery of evolutionary significant units. NOAA Technical Memorandum NMFS-NWFSC-42.

Pearse, D. E., C. M. Eckerman, F. J. Janzen, and J. C. Avise. 2001. A genetic analogue of 'mark–recapture' methods for estimating population size: an approach based on molecular parentage assessments. Molecular Ecology 10(11):2711-2718.

Rawding, D.J., C.S. Sharpe, and S. M. Blankenship. 2014. Genetic-based estimates of adult Chinook salmon spawner abundance from carcass surveys and juvenile out-migrant traps. Transactions of the American Fisheries Society 143(1):55-67.

Riester, M., P. F. Stadler, and K. Klemm. 2009. FRANz: reconstruction of wild multi-generation pedigrees. Bioinformatics 25(16):2134-2139.

Seamons, T. R., D. J. Rawding, P. Topping, and A. Bosworth. 2014. Spawner abundance estimates for Green River Chinook salmon. Washington Department of Fish and Wildlife, Olympia, WA.

Seamons, T. R., M. Crewson, J. Whitney, and P. Verhey. 2015. Genetic-based abundance estimates for Snohomish River Chinook salmon. Washington Department of Fish and Wildlife, Olympia, WA.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. Macmillan, New York.

Wang, J. 2004. Sibship reconstruction from genetic data with typing errors. Genetics 166(4):1963-1979.

Wang, J. 2012. Computationally efficient sibship and parentage assignment from multilocus marker data. Genetics 191(1):183-194.

Wang, J. 2013. An improvement on the maximum likelihood reconstruction of pedigrees from marker data. Heredity 111(2):165-174.

Wang, J., and A. W. Santure. 2009. Parentage and sibship inference from multilocus genotype data under polygamy. Genetics 181(4):1579-1594.

Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press, New York, NY.

APPENDIX

Appendix: Simulated population demography:

The following options were used in program SPIP to simulate the population:

[-A 4] The maximum age of any individual is 4 years, with no spawners of age 5 or older.

[--fem-prob-repro 0 0 .95 1.0] Females of age 3 have a 95% probability of spawning and 4 year olds have a 100% probability of spawning. This leads to an age distribution of spawners equal to 95%, and 5% for 3 and 4 year olds, respectively.

[--male-prob-repro 0 .1 .95 1.0] Males of age 2, 3, and 4 have a 10%, 95%, and 100% probability of spawning, respectively. This leads to an age distribution of spawners equal to 10%, 85.5% and 4.5% for 2, 3, and 4 year olds, respectively.

[--fem-postrep-die 1 1 1 1] All females die after spawning.

[--male-postrep-die 1 1 1 1] All males die after spawning.

[-f 0 0 .75 1] On average, females of age 3 produce only 75% as many offspring as do females of age 4.

[-m 0 .25 .75 1] On average, males of age 2 produce only 25% as many offspring as do males of age 3, and 3 year olds only produce 75% as many offspring as do males of age 4.

[--fem-rep-disp-par .25] This variable controls the variance in female reproductive success. The variance in female reproductive success was set at 25%, the ratio of the expected number of offspring to the variance in the number of offspring.

[--male-rep-disp-par .25] This variable controls the variance in male reproductive success. The variance in male reproductive success was set at 25%, the ratio of the expected number of offspring to the variance in the number of offspring.

[--mate-fidelity .5] Mating is not monogamous, but numbers of male mates per female are restricted. The selection of new mates by a female follows a Dirichlet process with a parameter of 0.5. The result is that typically each female will produce offspring with 2 or 3 different male mates, but rarely with more than 4 or 5.

CHAPTER 2

ESCAPEMENT ESTIMATES FOR COHO SALMON (*ONCORHYNCHUS KISUTCH*) USING TRANSGENERATIONAL GENETIC MARK-RECAPTURE

ABSTRACT

Coho salmon (*Oncorhynchus kisutch*) have declined dramatically in their southern range. Monitoring of coho salmon populations is needed to determine the status of current populations and mitigate future losses. Coho salmon often spawn in small, remote coastal streams where limited visibility and restricted access during spawning makes redd count estimates of total escapement inaccurate. A genetic parentage-based application of the traditional mark-recapture method (transgenerational genetic mark-recapture, tGMR) presents a theoretically high precision alternative for estimating coho salmon escapement. I applied tGMR to estimate total escapement of adult coho salmon in two northern California streams. In tGMR, the first sampling occasion is the collection of adult coho salmon tissue (typically from traps and/or carcasses) and the second sampling occasion is the collection of juveniles from out-migrant traps. Recaptures are determined by the number of parent-offspring relationships identified by genetic parentage analysis. Total escapement is estimated using estimators that have the form of a simple Lincoln-Petersen estimator. The approach is advantageous for highly fecund species because the recapture sample size within *C*, may exceed *N* so that the variance of tGMR estimators can potentially be much less than for conventional Lincoln-Peterson estimators. Applying this method to coho salmon in two northern California streams resulted in total escapement estimates using the “with replacement” and “without replacement” models of 576 and 444 (Mill Creek, 2011-2012), 131 and 193 (Mill Creek, 2012-2013), and 430 and 468 (Freshwater Creek, 2012-2013). Transgenerational mark-recapture may provide a more

precise and less invasive alternative for estimating total escapement of coho salmon in small coastal streams.

INTRODUCTION

Population census size (*N*) is one of the most important and difficult parameters to estimate (Luikart et al. 2010). The use of non-invasive genetic tagging approaches to estimate *N* within a traditional Lincoln-Peterson framework has become increasingly popular in fisheries and wildlife management across a wide range of species, including the humpback whale (*Megaptera novaeangliae*; Palsbøll et al. 1997), grizzly bear (*Ursus arcto;* Boulanger and McLellan 2001), northern hairy-nosed wombat (*Lasiorhinus krefft*i; Banks et al. 2003), wolverine (*Gulo gulo*; Mulders et al. 2007), black bear (*Ursus americanus*; Dreher et al. 2007), Chinook salmon (*Oncorhynchus tshawytscha*; Hamazaki and DeCovich 2014), and muskellunge (*Esox masquinongy;* Miller et al. 2015). In addition to those studies that have used genetic methods as a tagging approach, a growing number studies have applied a genetic parentage-based analysis within a mark-recapture setting (Jones and Avise 1997; Pearse et al. 2001; Fiumera et al. 2002; Rawding et al. 2014; Bravington et al. 2016). An appealing attribute of parentage-based estimates of abundance is that they can potentially provide variance estimates that are much less than conventional Lincoln-Peterson estimators.

Rawding et al. (2014) introduced a genetics-based version of mark-recapture that involves parentage analysis for estimating the total number of salmon returning to their freshwater spawning habitat (or escapement), termed transgenerational genetic mark-recapture (tGMR). In tGMR, the first sampling occasion is the collection of adult salmon tissue (typically from traps and/or carcasses) and the second sampling occasion is the

collection of juveniles from out-migrant traps. Recaptures are the number of parent-offspring relationships identified by genetic parentage analysis. An advantage of this approach results from the high fecundity of salmon, which results in tagging large numbers of offspring from a small number of parents (Anderson and Garza 2005). This can lead to a recapture sample size that exceeds the total population. Since variance estimation in mark-recapture is strongly influenced by recapture sample size, the variance of tGMR estimators can potentially be much less than conventional Lincoln-Peterson estimators or redd count estimates of total escapement. Additionally, tGMRis less invasive than approaches that involve handling of live adults as it may only require handling of adult carcasses following the completion of spawning and juveniles during outmigration. In contrast, conventional Lincoln-Peterson estimators often involve the use of weirs that can disrupt spawning migrations of salmon. The tGMR method has shown promise for Chinook salmon, therefore warranting further investigation into coastal populations of coho salmon of conservation concern in California (Rawding et al. 2014; Seamons et al. 2014; Seamons et al. 2015).

There are currently two versions of the tGMR method, termed the "binomial" or "with replacement" and "hypergeometric" or "without replacement" models (Rawding et al. 2014). In the binomial or "with replacement" method, total escapement is estimated as:

$$\widehat{N} = \frac{MC}{R}$$

where adult salmon sampled in the first occasion and successfully genotyped are considered the marks (*M*). Let *n* equal the number of juvenile out-migrants sampled and successfully genotyped on the second occasion. Each genotyped juvenile has the potential of being assigned to both a male and female parent in *M* through parentage analysis, thus representing two capture opportunities for each one of the juveniles genotyped. For this reason, $C = 2n$. A juvenile from the second sample could be assigned to no parents in *M*, one parent in *M* or two parents in *M*. The recaptures, *R*, are equal to the sum of the total parentage assignments (0, 1 or 2) for the "with replacement" model.

The hypergeometric approach was also presented in Rawding et al. (2014) and applies a "without replacement" sampling framework, which is also appropriate for the estimation of *N* (Seber 1982):

$$\widehat{N} = \frac{MC'}{R'}$$

In comparison to the "with replacement" model, the "without replacement" approach requires an estimate of the total number of distinct parents that gave rise to the juvenile subsample (*C'*). This is the total number of individuals from *M* that were assigned as parents of juveniles through genetic parentage analysis, *R*', plus the number of additional "unmarked" individuals inferred to have been parents by parentage analysis. In contrast with *R* where the total number of parent-offspring assignments is tallied, *R*' is a simple count of the number of individuals in *M* that were assigned as parents

Genetic approaches have become very popular within fisheries and wildlife management, but care must be taken in study design and implementation (Marrucco et al.

2011). Like all mark-recapture approaches, generating an unbiased escapement estimate using tGMR requires careful examination of model assumptions (Lukacs and Burnham 2005a). For example, parentage analysis must be accurate and, ideally, without error. Genotyping methods, laboratory procedures, and proper genetic analysis software can aide in meeting this assumption. Microsatellite DNA analysis has been used for individual genotyping, but this method can result in high error rates and raises concern for the accuracy of genetic mark-recapture approaches (Sethi et al. 2014). Alternatively, biallelic single nucleotide polymorphisms (SNPs) are more attractive in cases of individual identification and parentage testing because they exhibit low scoring error rates and fewer allelic drop outs in degraded samples (Morin et al. 2004). This study was the first to use high-throughput SNPs to conduct parent-offspring reconstruction and analysis in a tGMR framework.

I generated estimates of total adult escapement using tGMR for wild coho salmon (*Oncorhynchus kitsuch*) returning to Freshwater and Mill creeks, which are within the Southern Oregon Northern California Coast (SONCC) Evolutionarily Significant Unit. Coho salmon in the SONCC ESU are listed as threatened under both the Federal Endangered Species Act and California's Endangered Species Act (CESA; CDFG 2002). Spawner estimates are considered the single most important measurement needed for ESA listed salmon species (Crawford and Rumsey 2011) and the California Coastal Salmonid Monitoring Plan (CMP; Adams et al. 2011) highlights the goal of monitoring a stratified subset of coho salmon populations and generating abundance estimates for monitoring status and trends. For small coastal streams without life cycle monitoring

stations, new methods are needed for monitoring trends in abundance through time. The estimates of total adult escapement using tGMR for Freshwater Creek coho salmon were compared to traditional escapement estimates based on conventional Lincoln-Peterson mark-recapture methods. I discuss tGMR findings in the context of factors that may bias estimates resulting from the approach, including immigration/emigration, uncertainty of parameters estimated using genetic parentage analysis, and the reproductive success of an individual affecting capture probabilities.

METHODS

Study Sites

Mill Creek is a second-order tributary to the Smith River, Del Norte County, California. The Mill Creek watershed drains an area of 99.7 km^2 and enters the Smith River at Jedidiah Smith Redwoods State Park, and has two main spawning tributaries, East Fork Mill Creek (watershed area= 37 km^2) and West Branch Mill Creek (watershed area= 24 km^2). The Mill Creek watershed is characterized by steep, mountainous terrain typical of northern California Coast Ranges. Elevations range from 21-710 meters above mean sea level (Madej et al. 1986). From 1974 to 1981 the United States Geological Survey (USGS) monitored stream discharge for Mill Creek 1 km below the confluence of the East Fork and West Branch (Stillwater Sciences 2002). During this time, mean discharge of Mill Creek was ~3 cms. Most precipitation occurs between October and March with mean annual precipitation ranging from 152 – 381 cm.

Mill Creek is considered to have the primary coho salmon population within the Smith River watershed and is also inhabited by Chinook salmon (*Oncorhynchus tshawytscha*), chum salmon (*O. keta*), steelhead trout (*O. mykiss*), coastal cutthroat trout (*O. clarkii*), western brook lamprey (*Lampetra richardsoni*), Pacific Lamprey (*Entosphenus tridentatus*), prickly sculpin (*Cottus asper*), coastrange sculpin (*Cottus alueticus*), threespine stickleback (*Gasterosteus aculeatus*), Klamath smallscale sucker

(*Catostomus rimiculus*) and American shad (*Alosa sapidissima*) (Albro and Gray 2002; Justice 2007; McLeod and Howard 2010).

Freshwater Creek is a fourth-order tributary to Humboldt Bay via Eureka Slough in Humboldt County, California. Freshwater Creek watershed drains an area of 92.3 km^2 with elevation ranging from ~0 to 823 m (Ricker et al. 2014). The main-stem supports 14.5 km of anadromous fish habitat. There are five main tributaries to Freshwater Creek that each provide up to 4 km of anadromous fish access: Cloney Gulch, Graham Gulch, Little Freshwater, McCready Gulch, and South Fork Freshwater. The upper basin consists of rocky substrate and is managed for timber production by Humboldt Redwood Company. The lower basin is confined with levees and is dominated by fine sediments with residential and limited commercial development near the mouth. Annual rainfall averages ~125 cm across the watershed with the majority accumulating between October and April. During the peak of the rainy season, stream discharge ranges from 0.5 to > 57 cms.

In addition to coho salmon, Freshwater Creek is also inhabited by Chinook salmon, steelhead trout, coastal cutthroat trout, threespine stickleback, prickly sculpin, coastrange sculpin, Pacific giant salamanders, Pacific lamprey, and western brook lamprey.

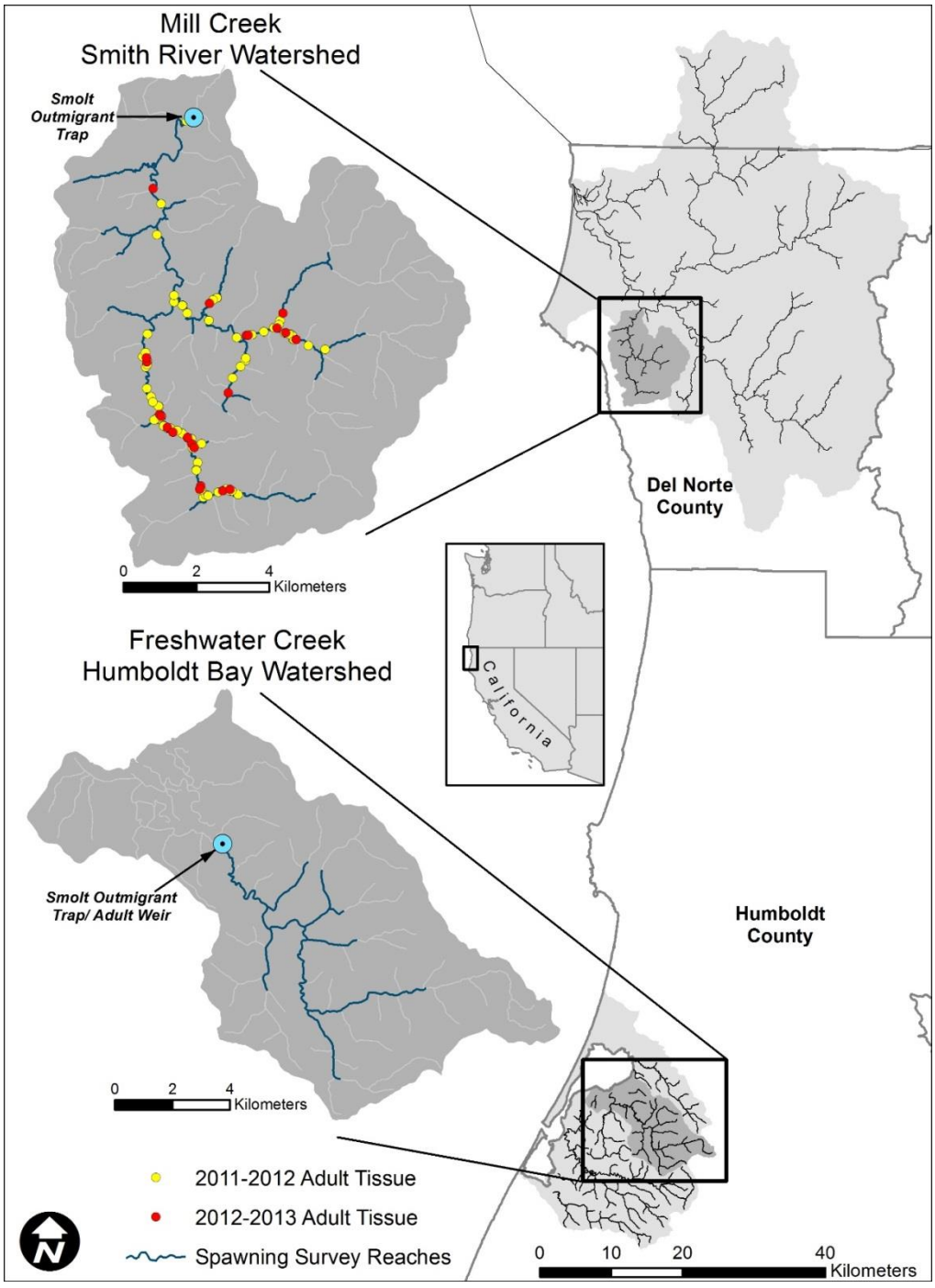


Figure 1. Location of Freshwater and Mill creek study locations in northern California, USA. Indicated are the locations of adult weir, carcasses recovery (red and yellow dots), and smolt outmigrant trap sites.

Field Data Collection

Handling and processing of all tissue samples used in this study was approved under the HSU IACUC permit number 13/14.F.122.A.

Adult Tissue Collection

Mill Creek – Adult coho salmon samples consisted of in-river collections of carcasses obtained from October through February during spawning surveys following protocols defined by Gallagher et al. (2007) and recommended by Adams et al. (2011) (Figure 1). Every deceased fish encountered during carcass surveys was identified to species, measured (fork length (cm)), examined for marks, and a tissue sample was collected for genetic analysis. During the 2011-2012 spawning season (October - February), 85 surveys were conducted and a total of 82 coho salmon carcasses were encountered with tissue collected from 63 carcasses. Carcasses ranged in fork length from 38 cm to 84 cm. During the 2012-2013 spawning season (November - February), 249 surveys were conducted and a total of 23 coho salmon carcasses were encountered with tissue collected from all 23 carcasses. Carcasses ranged in size from 39 cm to 72 cm. For more detailed explanation of survey protocols, see Garwood and Larson (2014).

Freshwater Creek – Adult coho salmon were collected during upstream migration as they entered the Humboldt Bay Fish Action Council weir (HFAC weir) (Figure 1). The weir is located approximately 8 river kilometers upstream from Humboldt Bay and operated continuously from the first fall rains in October/November until June, except during periods of high discharge when water levels render the weir inoperative. From November

11, 2012 to March 10, 2013, tissue was collected from 109 adult coho salmon at the HFAC weir. In addition, from November to June, 109 spawning ground surveys were conducted, 162 live coho salmon were observed and 50 carcasses were sampled. Each sampled fish was identified to species, measured for fork length (cm), examined for marks, and caudal fin clips collected for genetic analysis. Survey protocol followed Gallagher et al. (2007). Encountered adults ranged in size from 40 cm to 80 cm. For details regarding survey procedures see Ricker et al. (2014).

Juvenile Tissue Sampling

Juvenile tissues were collected during out-migration from March through June (about 4 months). On Mill Creek, modified pipe traps were installed in 2013 and used to collect juvenile coho salmon at the mouth of the West Branch (n_{WB}= 1,108) and the mouth of the East Fork (n_{EF}= 947). In 2014, a rotary screw trap was installed just below the confluence of the West Branch and East Fork of Mill Creek. Juveniles were sampled over a 15-week period (n_{MC}= 1,659). For details on trapping procedures and protocols, see Larson (2012) and Walkley et al. (2014). In 2014 on Freshwater Creek, tissues from juvenile coho salmon (n_{FW}= 2,409) were sampled using a trap box at the HFAC weir. For details on the trap and protocols see Anderson et al. (2015).

Juvenile Abundance Estimates and Tissue Subsampling

To ensure the genotyped juveniles represented the entire out-migrating population, an estimate of the entire smolt out-migrating population was generated and juvenile tissues were proportionally sampled. Estimates of coho salmon smolt abundance were generated using a single trap mark-recapture strategy following McLeod and

Howard (2010) and Ricker and Anderson (2011). Each day, a representative sample of previously unmarked coho salmon smolts were tagged with individually numbered PIT tags (Prentice et al. 1994) and given a fin clip. During periods of high smolt abundance, additional coho salmon smolts were marked only with fin clips. Following tagging and/or marking, fish were held in flow-through live cars to check for handling/marking mortality before being released upstream of the trap. Newly marked fish were released one to three pool-riffle sequences upstream of the trap. Mark-recapture of fin clips was broken into weekly strata and estimates of weekly abundance were calculated using Darroch Analysis of Ranked Regression (DARR 2.0.2; Bjorkstedt 2005; Bjorkstedt 2010) in R (R development Core Team 2013). Weekly abundance estimates were then used to estimate the total out-migrating population of coho salmon smolts. In 2013, total abundance of out-migrating coho salmon smolts were estimated at $\widehat{N}_{WB} = 6157$ and $\widehat{N}_{EF} = 3762$ for the West Branch and East Fork of Mill creek, respectively. These results were then combined ($\widehat{N}_{WB} + \widehat{N}_{EF}$) to give a single estimate of $\widehat{N}_{MC}$ =9919. In 2014, total abundance of out-migrating coho salmon smolts was estimated at $\widehat{N}_{MC} = 9956$, and $\widehat{N}_{FW} = 15724$, for Mill and Freshwater creeks, respectively.

To ensure representative sub-sampling of tissue for the entire smolt out-migration period, weekly abundance estimates were divided by the total abundance estimate to give weekly proportions of out-migrating smolts. These weekly proportions were multiplied by the total desired number of tissue samples for genotyping, yielding the number of tissues to sub-sample from the weekly strata. Available tissue samples from each week were organized in ascending order, first by date and then by sample number. Tissues

were sub-sampled systematically from each weekly stratum after a random start. If tissue samples were needed from a weekly stratum without tissue samples available, the number of samples needed was pooled with the previous weekly stratum ensuring that a representative number of samples was collected. For the 2013 juvenile trapping season, estimates for West Branch and East Fork trap sites were pooled ($\widehat{N}_{MC}$ =9919) and then sub-sampled proportionately (n_{WB}= 310, n_{EF}= 190) as described above to achieve the desired number of samples (n_{MC}= 500) (Table 1). For the 2014 outmigration season, juveniles were sampled and genotyped from Mill Creek (n_{MC}= 501) (Table 2), and Freshwater Creek (n_{FW}= 1,002).

Table 1. Total recruitment estimated for Mill Creek by DARRv2 over the 12 week trapping period in 2013 for West Branch (WB, $\widehat{N}_{WB}$= 6156.83) and 11 week trapping period for the East Fork (EF, $\widehat{N}_{EF}$= 3761.87). The pooled smolt population ($\widehat{N}_{WB} + \widehat{N}_{EF}$) estimate for the trapping season was $\widehat{N}_{Mill}$= 9918.70. $\widehat{N}_{WB_i}$ are the weekly estimates for West Branch Mill Creek. $\widehat{N}_{WB_i}/\widehat{N}_{WB}$ are the estimated weekly proportions West Branch Mill Creek. $\widehat{N}_{EF_i}$ are the weekly estimates for East Fork Mill Creek. $\widehat{N}_{EF_i}/\widehat{N}_{EF}$ are the estimated weekly proportions for East Fork Mill Creek. Samples genotyped are the number of samples originally genotyped for each trapping week. Samples collected are the total number of samples collected for each week during the trapping season.

	West Branch				East Fork			
Trapping Week, i	$\widehat{N}_{WB_i}$	$\widehat{N}_{WB_i}/\widehat{N}_{WB}$	Samples Genotyped	Samples Collected	$\widehat{N}_{EF_i}$	$\widehat{N}_{EF_i}/\widehat{N}_{EF}$	Samples Genotyped	Samples Collected
1	25	0.0041	1	3	11.80	0.0031	1	2
2	155	0.0252	8	7	48.50	0.0129	2	15
3	270	0.0439	14	14	3.93	0.0010	0	2
4	47.74	0.0078	2	8	44.56	0.0118	2	15
5	229.87	0.0373	12	81	505.93	0.1345	26	133
6	552.14	0.0897	28	138	1207.17	0.3209	61	181
7	1747.06	0.2838	88	181	1341.13	0.3565	68	259
8	2153.14	0.3497	109	292	351.36	0.0934	18	191
9	610.90	0.0992	31	200	123.68	0.0329	6	83
10	195.50	0.0318	10	99	42.56	0.0113	2	28
11	59.43	0.0097	3	39	81.25	0.0216	4	38
12	111.05	0.0180	6	46	*NA*	*NA*	*NA*	*NA*

Table 2. Total recruitment estimated by DARRv2 over the 15 week trapping period in 2014 for Freshwater Creek was $\widehat{N}_{FW}$= 15724.17 and $\widehat{N}_{MC}$= 9956.34 for Mill Creek. $\widehat{N}_{FW_i}$ are the weekly estimates for Freshwater Creek. $\widehat{N}_{FW_i}/\widehat{N}_{FW}$ are the estimated weekly proportions for Freshwater Creek. $\widehat{N}_{MC_i}$ are the weekly estimates for Mill Creek. $\widehat{N}_{MC_i}/\widehat{N}_{MC}$ are the estimated weekly proportions for Mill Creek. Samples genotyped are the number of samples originally genotyped for each trapping week. Samples collected are the total number of samples collected for each week during the trapping season.

	Mill Creek				Freshwater Creek			
Trapping Week	$\widehat{N}_{FW_i}$	$\widehat{N}_{FW_i}/\widehat{N}_{FW}$	Samples Genotyped	Samples Collected	$\widehat{N}_{MC_i}$	$\widehat{N}_{MC_i}/\widehat{N}_{MC}$	Samples Genotyped	Samples Collected
1	22.51	0.0014	1	1	94.06	0.0094	5	10
2	167.18	0.0106	11	52	432.67	0.0435	22	47
3	771.60	0.0491	49	111	366.99	0.0369	18	50
4	106.10	0.0067	7	32	1258.27	0.1264	63	80
5	1092.97	0.0695	70	281	1433.03	0.1439	72	39
6	2363.97	0.1503	150	412	657.90	0.0661	33	150
7	2814.49	0.1790	179	282	1155.63	0.1161	58	140
8	2681.23	0.1705	171	289	1293.12	0.1299	65	303
9	1641.75	0.1044	104	183	771.81	0.0775	39	102
10	684.31	0.0435	44	190	1124.07	0.1129	56	380
11	1210.84	0.0770	77	263	803.92	0.0807	40	222
12	999.22	0.0635	64	237	415.59	0.0417	21	94
13	967.25	0.0615	62	78	133.57	0.0134	7	51
14	155.13	0.0099	10	0	15.71	0.0016	1	7
15	45.63	0.0029	3	0	0.00	0.0000	0	0

Genetic Data Collection

Molecular Methods

DNA was extracted from dried fin clips using the DNeasy 96 filter-based nucleic acid extraction system on a BioRobot 3000 (Qiagen, Inc.), following the manufacturer's protocols. Extracted DNA was diluted 2:1 with distilled water and used for assay of 96 SNPs. Genotyping was executed with a standardized set of 96 SNP markers developed and validated for California coho salmon populations (Smith et al. 2006; Campbell and Narum 2011; Starks et al. 2015) including a locus added for distinguishing Chinook salmon from coho salmon (Starks et al. 2016) (Supplementary material, Table 1). Genotyping was performed at the NOAA Southwest Fisheries Science Center (Santa Cruz, California) using a Fluidigm EP1 real-time PCR instrument. This platform uses 96.96 Fluidigm arrays that evaluate 96 samples at 96 loci in parallel. A single 96.96 array results in 9,216 SNP genotypes, with each sample genotyped for the same set of 96 SNP loci.

Data were filtered to include only those individuals with at least 81 genotyped loci. Tests for conformance to Hardy–Weinberg proportions for each locus in each adult collection were conducted using the Markov Chain method (dememorization number 1000, 100 batches, 1000 iterations per batch) and observed heterozygosity was generated in Genepop 4.5.1 (Raymond and Rousset 1995; Rousset 2008) (Supplementary material, Table 1). Expected heterozygosity was estimated using the software Genotype Viewer (2007).

Parentage Analysis

Pedigrees were reconstructed using the maximum likelihood algorithms (Wang 2004; Wang and Santure 2009; Wang 2012; Wang 2013) implemented in the software COLONY (version 2.0.6.1) (Jones and Wang 2009). COLONY uses multilocus genotypes to infer sibship and parentage among individuals using a full-pedigree likelihood method. Individuals were divided into an offspring sample, a candidate father sample, and a candidate mother sample. COLONY reconstructs full- and half-sibling family clusters, parent-offspring pairs including inference of parents that were not in the original sample of candidate fathers or mothers, which provides the necessary information to estimate total escapement using tGMR statistical methods. One limitation is that computation time is demanding. Analysis of a single run can range from minutes to days to complete depending on initial parameters and sample sizes. For example, all three iterations for Freshwater Creek took about 20 hours to complete (Intel i5 dual core processor with 8 GB of RAM). The settings for parentage analysis using COLONY were: a polygamous mating system, no inbreeding or clones, species are dioecious and diploid, length of run and likelihood precision are set to medium, full likelihood analysis, no sibship priors, genotyping error rates at 0.0001 per marker, and unknown allele frequency. COLONY was run three times for each dataset with identical settings but with different random number seeds. Results of the three runs were compared to evaluate convergence of the point estimate, defined with a coefficient of variation (CV) of less than 5% (Seamons et al. 2014) across the three runs for the “with replacement” and

"without replacement" models. When convergence was reached, the first estimate of the three runs was reported.

In several cases, the sex of candidate parents used in parentage analysis was unknown, including Mill Creek 2011 (N= 7 parents of unknown sex), Mill Creek 2012 (N= 1), and Freshwater Creek 2012 (N= 15). Individuals of unknown sex were entered as both candidate fathers and candidate mothers in the parentage analysis, enabling COLONY to potentially determine sex by matching an individual of unknown sex with a breeding partner of known sex.

Coho salmon juveniles can demonstrate a two-year freshwater life history (Bell and Duffy 2007), and inclusion of these individuals may bias escapement estimates generated by tGMR. To determine the extent to which two-year freshwater life history individuals may bias tGMR estimates herein, juveniles exhibiting the two-year freshwater life history were identified by parentage analysis for Mill and Freshwater creeks using candidate parent genotypes from 2011-2012 and offspring samples from 2014. While this approach can provide an indication of the frequency of the two-year freshwater life history in the study areas, only individuals that had at least one parent in the adult sample can be identified through parentage analysis. Any individuals exhibiting a two-year freshwater life history were excluded from the 2012-2013 tGMR escapement estimates because they are not in the same cohort of the 2014 offspring sample.

Estimators

Escapement estimates were generated using two versions of tGMR, termed the "with replacement" and "without replacement" models (Rawding et al. 2014). In the "with replacement" model, total escapement is estimated as:

$$\widehat{N}_{WR} = \frac{MC}{R}$$

where adult salmon sampled in the first occasion and successfully genotyped are considered an individual mark (*M*). Let *n* = the number of juvenile out-migrants sampled and successfully genotyped on the second occasion. Each genotyped juvenile has the potential of being assigned to both a male and female parent in *M* through parentage analysis, thus representing two capture opportunities for each one of the juveniles genotyped. For this reason, $C = 2n$. Recaptures or *R* is the number of times adults present in *M* are detected in the juvenile sample. Conditional variance of the "with replacement" estimator was estimated as (Ricker 1975):

$$\widehat{V}_{WR} = \frac{M^2 C(C-R)}{R^3}$$

The hypergeometric approach was also presented in Rawding et al. (2014) and applies a "without replacement" sampling framework, which is also appropriate for the estimation of *N* (Seber 1982):

$$\widehat{N}_{WOR} = \frac{MC'}{R'}$$

where *C'* is the total number of distinct parents that gave rise to the juvenile subsample, and the recaptures, or *R'* is the number of distinct parents from the first sampling occasion (*M*) that are assigned as parents of individuals from the juvenile genotypes in the second sampling occasion. Conditional variance of the "without replacement" estimator was estimated as (Williams et al. 2002):

$$\hat{V}_{WOR} = \frac{MC'(C' - R')(M - R')}{R^3}$$

RESULTS

Carcass tissue samples were collected from Mill Creek during spawning years 2011-2012 (n= 63) and 2012-2013 (n= 23) and 35 and 15 individuals were successfully genotyped, respectively. During the 2013 out-migrant season juveniles were proportionally subsampled (n= 500) and 494 were successfully genotyped. During the 2014 out-migrant season juveniles were proportionally subsampled (n= 501), and 485 were successfully genotyped. Adult tissue samples were collected from Freshwater Creek during the 2011-2012 (n= 133) and 2012-2013 (n= 107) spawning years, and 132 individuals and 106 individuals were successfully genotyped, respectively. Of the 1,002 juvenile out-migrants proportionally subsampled from the 2014 out-migrant season, 936 were successfully genotyped. Eighteen of the 2014 Freshwater Creek juveniles were assigned to parents from the 2011-12 spawning year, indicating that they were two years old, and therefore removed from the estimates for 2012-13 spawning year. After correcting for juveniles exhibiting the two year freshwater life history, and including adults of unknown sex as both candidate males or females, the following number of individuals for each site and year estimate were used for parentage analysis with COLONY; Mill Creek 2011-2012 (offspring sample= 494, candidate father sample= 19, candidate mother sample= 23), Mill Creek 2012-2013 (offspring sample= 485, candidate father sample= 5, candidate mother sample= 11), and Freshwater Creek 2012-2013 (offspring sample= 928, candidate father sample= 71, candidate mother sample= 50).

Genetic Analysis

Tests for conformance to Hardy-Weinberg proportions in each adult collection (Mill Creek 2011-2012, Mill Creek 2012-2013, Freshwater Creek 2011-2012, and Freshwater Creek 2012-2013) were non-significant ($\alpha = 0.05$) in all cases when corrected for multiple tests (Supplementary material, Table 1). The percent of polymorphic loci ranged from 92% to 98% (mean= 96%) among the four adult collections (Table 3). The species identification locus (Oki120255-113) confirmed that all tissue samples, including both juvenile and adult samples, were from coho salmon. Observed heterozygosity for the SNP loci ranged from 0.01 to 0.60 (mean= 0.32) for adults in all four populations (Supplementary material, Table 1). Expected heterozygosity for SNP loci ranging from 0.01 to 0.52 (mean= 0.33) for adults in all four populations (Table 3). Minor allele frequency (MAF) for polymorphic loci ranged from 0.01 to 0.50 (mean= 0.25) in all four years of adult collection.

Table 3. The total number of adult tissue samples collected and submitted for genotyping (Tissue Collected), number of adult tissue samples successfully genotyped at ≥81 loci (Successfully Genotyped), type of tissue submitted for genotyping (carcass or fresh), proportion of polymorphic loci (P), Hardy-Weinburg expected heterozygosity (H_e), and observed heterozygosity (H_o) for four adult collections of coho salmon from northern California.

	Tissue Collected	Successfully Genotyped	Tissue Type	P	H_e	H_o
Mill Creek 2011-2012	63	35	carcass	0.96	0.33	0.32
Mill Creek 2012-2013	23	15	carcass	0.92	0.34	0.32
Freshwater Creek 2011-2012	133	132	fresh/carcass[1]	0.98	0.33	0.32
Freshwater Creek 2012-2013	107	106	fresh/carcass[2]	0.98	0.33	0.32

[1]Tissue samples were collected from 128 live individuals and four carcasses.
[2]Tissue samples were collected from 101 live individuals and six carcasses.

Parentage Analysis

Mill Creek 2011-2012 – Of the 35 successfully genotyped adults, at least one offspring was assigned to 19 adults, including 10 of 16 females, seven of 12 males, and two of the seven adults of unknown sex. Two of the individuals of unknown sex were assigned as males. Five adults of unknown sex were not identified as parents of any individuals in the juvenile subsample, and therefore were not assigned a sex. In addition to the 19 individuals in the adult sample identified as parents, another 222 parents (113 males, 109 females) were inferred to have existed and to have been a parent for at least one of the genotyped juveniles in the parentage analysis. The number of offspring in the juvenile sample assigned to genotyped adult parents ranged from 1 to 12 (mean= 3.16, variance= 10.25), and the number of offspring assigned to inferred, unsampled parents ranged from 1 to 13 (mean= 4.18, variance= 7.09). The distribution of the number of offspring assigned to sampled parents (i.e., at least one offspring assigned) was not significantly different from that assigned to inferred, unsampled parents based on a permutation test (P= 0.12, Figure 2).

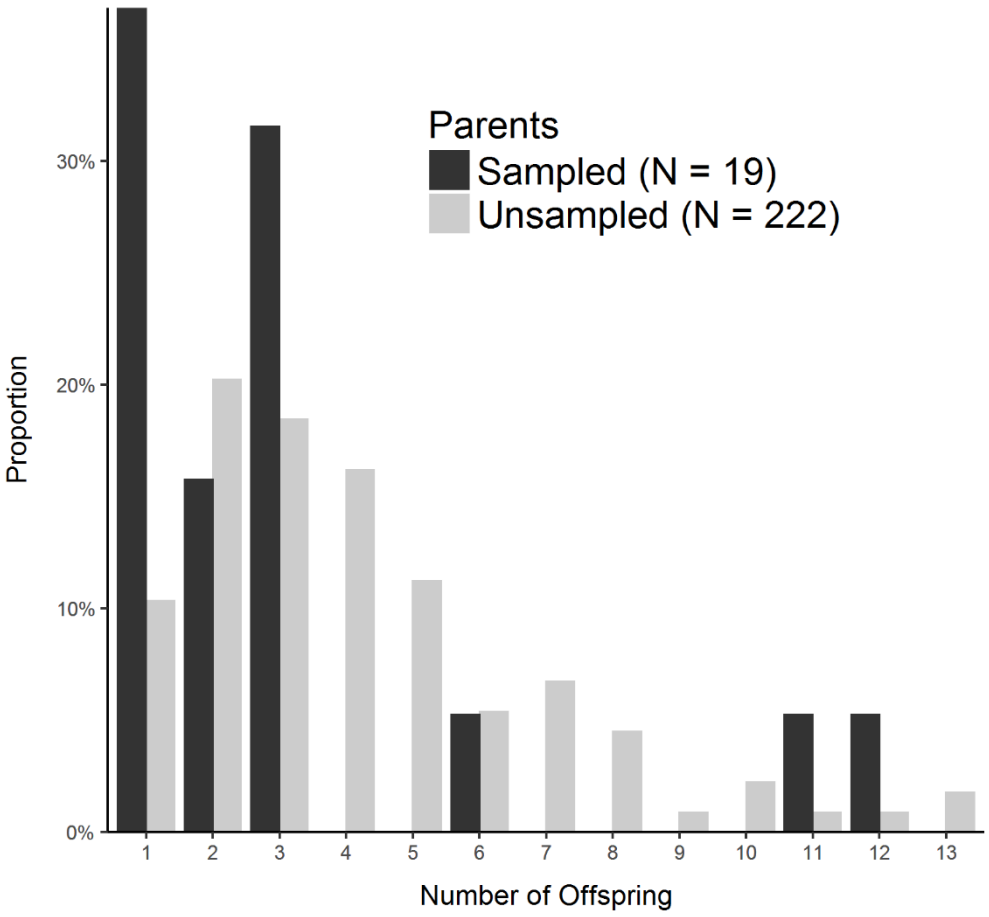


Figure 2. Frequency distribution of the number of offspring per parent for sampled crcasses and inferred, unsampled parents based on parentage analysis using COLONY for Mill Creek (2011-2012). The distributions were not significantly different, based on a permutation test (P = 0.12).

Mill Creek 2012-2013 – Of the 15 successfully genotyped adults, at least one offspring was assigned to 14 adults, including nine of the 10 females, and all four males. A single adult of unknown sex was assigned as a male parent. In addition to the 14 individuals in the adult sample identified as parents, another 166 parents (88 males, 78 females) were inferred by COLONY to have been a parent for at least one of the genotyped juveniles. The number of offspring in the juvenile subsample produced by the genotyped adults ranged from 1 to 23 (mean= 7.93, variance= 37.61), and the number of offspring per inferred, unsampled parent ranged from 1 to 19 (mean= 5.17, variance= 14.33). The distribution of the number of offspring assigned to sampled parents (i.e., at least one offspring assigned) was significantly different from that assigned to inferred, unsampled parents based on a permutation test (P= 0.01, Figure 3). This may suggest that carcass sample size was not representative of the entire adult population, likely owing to its small size (N= 15), or that COLONY incorrectly reconstructed inferred parents.

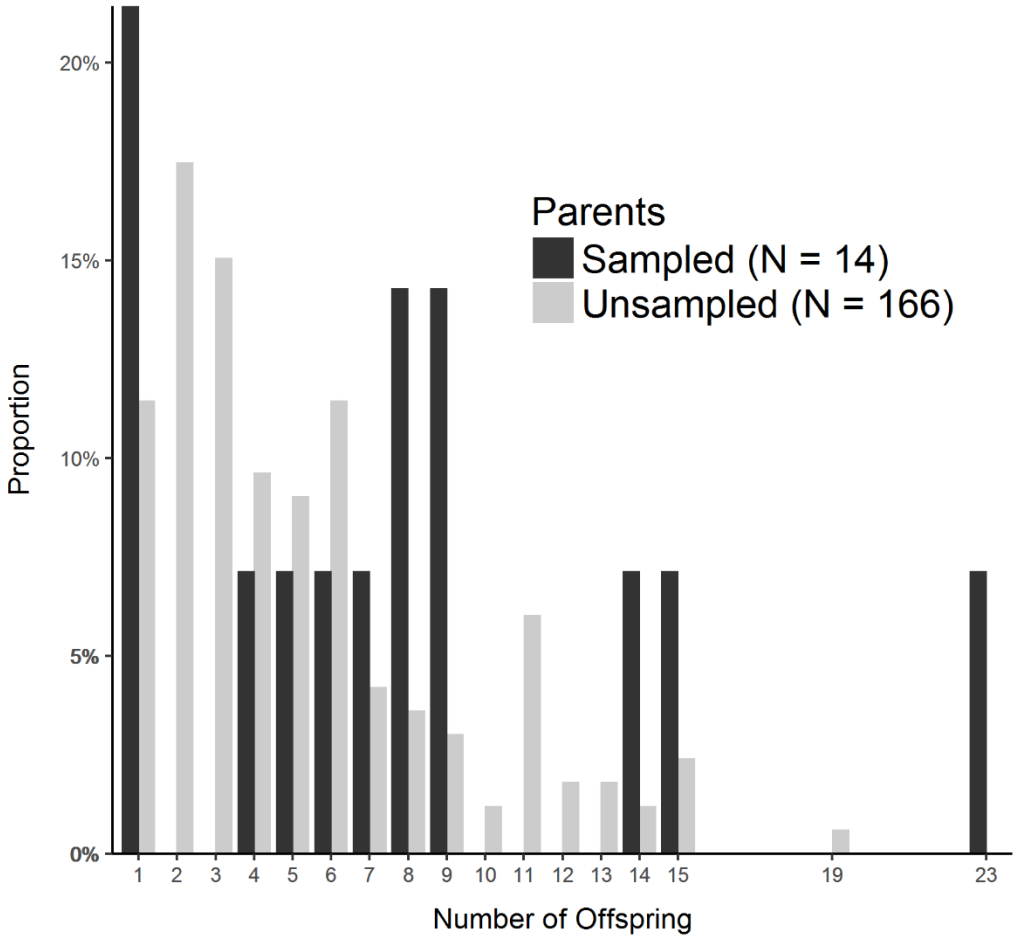


Figure 3. Frequency distribution of the number of offspring per parent for sampled carcasses and inferred, unsampled parents based on parentage analysis using COLONY for Mill Creek (2012-2013). The distributions were significantly different, based on a permutation test (P = 0.01).

Freshwater Creek 2012-2013 – Of 106 successfully genotyped adults, at least one offspring was assigned to 73 adults, including 25 of 35 females, 37 of 56 males, and 11 of 15 unknown sexes. Nine of the unknown sexes were assigned as females and two as males, while four did not produce any offspring in the subsample, and so were not assigned a sex. In addition to the 73 parents that COLONY assigned at least one offspring, another 249 parents (123 males, 126 females) were inferred by COLONY to have been a parent for at least one of the genotyped juveniles. The number of offspring per genotyped adults ranged from 1 to 29 (mean= 6.27, variance= 34.06), and the number of offspring per inferred, unsampled parent ranged from 1 to 31 (mean= 5.61, variance= 35.08). The distribution of the number of offspring assigned to sampled parents (i.e., at least one offspring assigned) was not significantly different from inferred, unsampled parents based on a permutation test (P= 0.41, Figure 4).

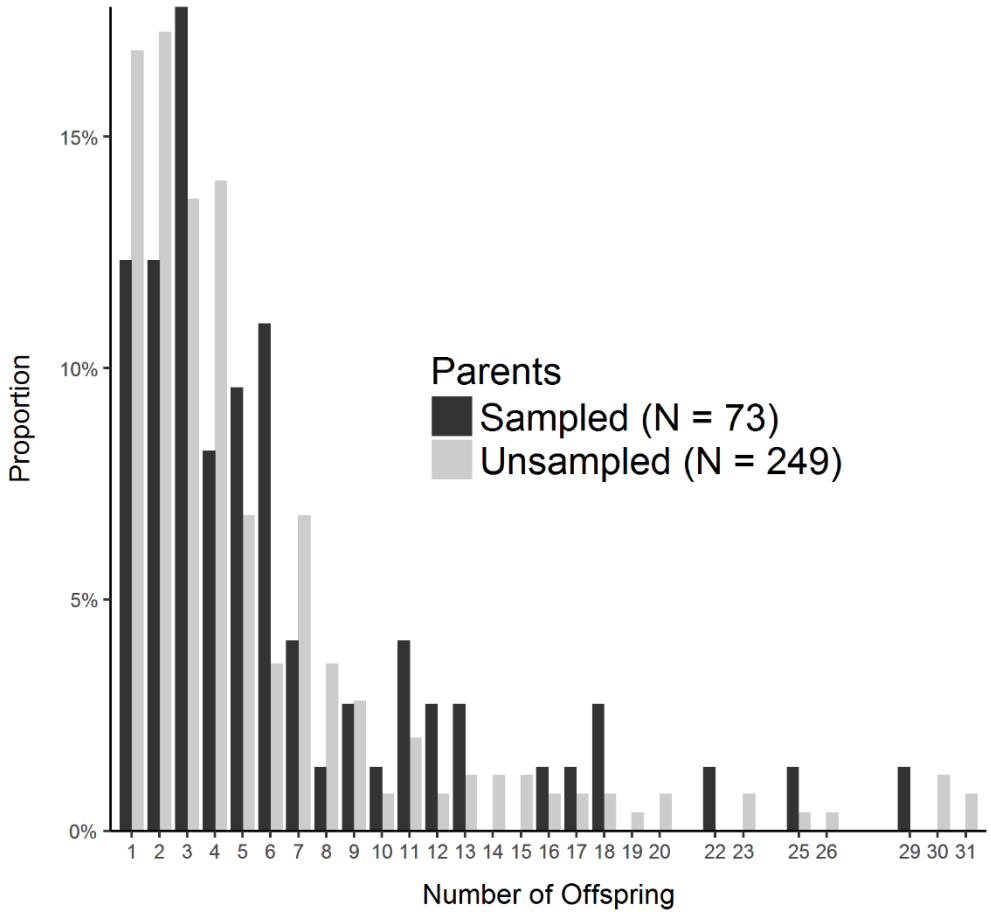


Figure 4. Frequency distribution of the number of offspring per parent for sampled carcasses and inferred, unsampled parents based on parentage analysis using COLONY for Freshwater Creek (2012-2013). The distributions were not significantly different, based on a permutation test (P = 0.41).

Escapement Estimates

Mill Creek 2011-2012 – A total of 35 adults and 494 juveniles were genotyped. Parentage analysis assigned 60 juveniles to at least one sampled parent (*R*) and estimated that 241 adults gave rise to the juvenile sample (*C*'). Nineteen of the genotyped adults had at least one offspring in the juvenile sample (*R*'). Convergence of the point estimates was met across all runs for the "with replacement" (CV= 0.019) and "without replacement" (CV= 0.024) models. The total escapement estimate using the "with replacement" model was 576 adults (95% CI= 509 – 644) (Table 4). The "without replacement" model estimated total escapement as 444 adults (95% CI= 312 – 576) (Table 5).

Mill Creek 2012-2013 – A total of 15 adults and 485 juveniles were genotyped. Parentage analysis assigned 111 juveniles to at least one sampled parent (*R*) and estimated that 180 adults gave rise to the juvenile sample (*C'*). Fourteen of the genotyped adults had at least one offspring in the juvenile sample (*R'*). Convergence of the point estimates was met across all runs for the "with replacement" (CV= 0.000) and "without replacement" (CV= 0.021) models. The total escapement estimate using the "with replacement" model was 131 adults (95% CI= 120 – 142) (Table 4). The "without replacement" model estimated total escapement as 193 adults (95% CI= 167 – 218) (Table 5).

Freshwater Creek 2012-2013 – A total of 106 adults and 928 juveniles were genotyped. Parentage analysis assigned 485 juveniles to at least one sampled parent (*R*)

and estimated that 322 adults gave rise to the juvenile sample (*C*'). Seventy three of the genotyped adults had at least one offspring in the juvenile sample (*R*'). Convergence of the point estimates was met across all runs for the "with replacement" (CV= 0.001) and "without replacement" (CV= 0.005) models. The total escapement estimate using the "with replacement" model was 430 adults (95% CI= 415 – 444) (Table 4). The "without replacement" model estimated a total escapement of 468 adults (95% CI= 414 – 521) (Table 5).

Table 4. Escapement estimates ($\widehat{N}_{WR}$) using the "with replacement" transgenerational genetic mark-recapture model for coho salmon from Mill and Freshwater creeks, where M is the marked individuals from the first sampling event, C is the captures from the second sampling event (or 2n, where n is the juvenile sample size), and R is the recaptures detected in C.

"With Replacement" model

	M	C	R	$\widehat{N}$	95% CI	CV
Mill Creek (2011-12)	35	988	60	576	509-644	0.0586
Mill Creek (2012-13)	15	970	111	131	120-142	0.0410
Freshwater Creek (2012-13)	106	1856	458	430	415-444	0.0166

Table 5. Escapement estimates ($\widehat{N}_{WO}$) using the "without replacement" transgenerational genetic mark-recapture model for coho salmon from Mill and Freshwater creeks, where M are the marked individuals from the first sampling event, C' are the captures from the second sampling event, and R' are the recaptures detected in C'.

"Without Replacement" model

	M	C'	R'	$\widehat{N}$	95% CI	CV
Mill Creek (2011-12)	35	241	19	444	312-576	0.1489
Mill Creek (2012-13)	15	180	14	193	167-218	0.0663
Freshwater Creek (2012-13)	106	322	73	468	414-521	0.0574

DISCUSSION

I estimated total escapement for natural populations of coho salmon in Mill and Freshwater creeks using tGMR methods. All total escapement estimates for "with replacement" and "without replacement" models met precision standards set by McElhany et al. (2000) with coefficients of variation (CV) of less than 15% (Table 4, Table 5). Estimates using tGMR were compared with traditional escapement estimates produced in Freshwater Creek, where total escapement is estimated by marking adults as they pass through the HFAC weir and examining carcasses for marks during spawning ground surveys. The Lincoln–Peterson mark-recapture model is then used to produce escapement estimates (Anderson et al. 2015). In the 2012-2013 spawning year the traditional mark-recapture method estimated 318 adult coho salmon, which is less than the tGMR estimates of 430 and 468 from the "with replacement" and "without replacement" models, but neither were significantly different from the traditional mark-recapture method using a t-test ($P > 0.05$). There are no escapement estimates based upon traditional mark-recapture methods for Mill Creek coho salmon available for comparison to tGMR estimates.

In all estimates of total escapement within this study, two-year old males were considered adults and included. Two-year old males may have offspring present in the juvenile subsample, thus eliminating them from the analysis can violate tGMR model assumptions. The California Department of Fish and Wildlife's protocol calls for inclusion of two year old males in total escapement estimates for ESA listed coho

salmon. Further, two year old males play a vital role in gene flow among brood years (Smith et al. 2014). Two year old males are typically excluded from escapement estimates in commercially or recreationally salmonid monitoring contexts.

Comparison of tGMR and Traditional Abundance Estimates

When comparing methods previously used on the Coweeman River (i.e., redd-based, AUC, and Jolly-Seber), Rawding et al. (2014) found tGMR estimates were in accordance with but slightly exceeded non-genetic approaches. Escapement estimates using tGMR for Chinook salmon on the Green River (Seamons et al. 2014) and Snohomish River (Seamons et al. 2015) were also larger than the traditional methods (i.e., redd-based). When comparing tGMR to redd-based counts, discrepancies among estimates may be due to the variables being estimated. For example, in redd-based estimates the number of redds constructed by females is estimated and then expanded to include the number of males present, after which live fish counts (those fish not seen on redds) are added to the estimate. More often in California, the number of redds are estimated and then multiplied by a fish per redd correction factor. As a result, this estimates the number of adults that successfully occupy redds and not the total escapement. In most redd-based estimates, pre-spawn mortality (Heard 1991, Quinn 2005) is not accounted for and may have caused the large differences in estimates for the Green and Snohomish River estimates where female carcasses full of eggs were sampled (Seamons et al. 2013). Redd misidentification issues can also biased redd-based estimates through redd superimposition, overlap in spawn timing (i.e., other species) (Gallagher

and Gallagher 2005), technician experience (Muhlfeld et al. 2006), and variation in redd characteristics (Crisp and Carling 1989). While one or more of these issues may have affected previous observations of differences between tGMR and traditional methods, the estimates for Freshwater Creek were based on a live fish mark-recapture approach and did not use redd-based counts for estimating total escapement.

Variance Estimation

Rawding et al. (2014) applied conventional Lincoln-Peterson variance estimators in a tGMR context, however these estimators are not appropriate for construction of confidence intervals using tGMR. For the "with replacement" estimator, an appropriate variance expression would need to account for the variation in the conditional expectation of estimator. The conditional expectation of this estimator varies according to the reproductive success of the specific selected adults genotyped on the first sampling occasion (M). Thus, an unconditional estimator of variance is needed to appropriately address realistic estimates of uncertainty. Currently, an unconditional variance estimator appropriate for the "with replacement" estimator does not exist (Mohr and Hankin pers. comm.).

For the "without replacement" estimator a considerable amount of bias and uncertainty may result from errors of estimation of model parameters (especially C', see Chapter 1, *Results*). The conditional variance estimator apparently used by Rawding et al. (2014) assumes that estimated model parameters (C', R') are unbiased and measured without error. An appropriate estimator for calculation of error of estimation for the

"without replacement" estimator would need to account for uncertainty in model parameters estimated from COLONY. No such variance estimator is yet available (Mohr and Hankin pers. comm.).

Adult Detection

The proportion of marked adults detected through juvenile genotypes ranged from 54% to 93%. Similar results were obtained for the simulated populations in chapter 1 (Table 4) where on average, the proportion of marked adults detected through juvenile genotypes ranged from 59% to 89%. As the juvenile sample size increased, so did the proportion of adults detected. This proportion is also comparable to previous studies. For example, Rawding et al. (2014) detected 83% of the marked adults through the juvenile sample. As a result, there is a high probability that an adult present in *M* will be detected in the juvenile sample.

Carcass Samples

In Mill Creek, where all adult samples consisted of carcasses, tissue degradation constrained successful genotyping and rendered a large number of adult samples unusable for parentage analysis (see Baumsteiger and Kerby 2009; Copeland et al. 2009). Tissue degradation associated with use of carcasses is hypothesized to have resulted in 35% to 44% of the Mill Creek adult samples being unsuitable for analysis (Table 3). In contrast, 99% of the samples collected from live adults from Freshwater Creek yielded at least 81 loci genotyped. Also, all juvenile tissue was collected from live individuals

resulting in 93% to 99% of the samples being of use in this tGMR study. With the collection of carcass tissue being limiting with only 20% of the adult population (Seth Ricker, California Department of Fish and Wildlife, pers. comm.), increases in carcass tissue collection effort is recommended. This may improve total escapement estimates by increasing the adult sample size.

tGMR Assumptions

For tGMR total escapement estimates to be approximately unbiased, assumptions must be met. If all assumptions are not met, then tGMR estimators may be biased and conditional variance estimators used in this report will underestimate uncertainty in estimation of escapement. The following assumptions are adjustments to the relevant simple Lincoln-Peterson estimator assumptions (Seber 1982) as seen appropriate for the tGMR context.

Assumption 1: No emigration or immigration between the two sampling events.

Emigration in tGMR occurs when a juvenile belonging to one of the parent samples in the first occasion is not available for capture during the second sample occasion. For coho salmon, the emigration assumption may be violated as coho salmon smolts may emigrate from natal streams prior to the initiation of out-migrant field collections (Rebenack 2015). It is unknown how many juveniles emigrate before or after trapping season.

Immigration occurs in two obvious instances when estimating escapement of coho salmon using tGMR. First, juvenile coho salmon may rear in freshwater for two years

before out-migrating to the ocean (Bell and Duffy 2007). These two-year-old smolts would be immigrants from the previous cohort and not belong to the cohort under study. Further, these two-year-old smolts would be considered emigrants with respect to their birth cohort. Herein, parentage analysis was conducted to identify juveniles that exhibited the two-year freshwater life history. In Mill Creek, none of the genotyped juveniles were found to be two years old, but in Freshwater Creek 18 juveniles were found to be two years old. These 18 individuals were removed from the analysis because inflating the number of captures would result in downward bias of $\widehat{N}$. These individuals were detected using parentage analysis of adults from two years before the offspring sample. Using parentage analysis, juveniles exhibiting a two year freshwater life history can only be detected if they had at least one parent in the genotyped adult sample. Thus, it is unlikely that all two year old juveniles were removed from this analysis. An alternative approach would involve aging all juveniles (i.e., scale analysis), and removing those individuals who did not belong to the cohort under study. However, due to the possibility of interpretation error, it is unclear whether scale analysis can be used to confidently identify individuals exhibiting a two-year freshwater life-history (Maceina et al. 2007). Second, immigration may result when juvenile coho salmon enter from a nearby stream for over-winter rearing and are therefore not progeny of adults genotyped in the first sampling occasion. Juvenile coho salmon have been documented to immigrate from natal tributaries and move upstream in the mainstems of rivers (Hance et al. 2016) and such movement is likely in Freshwater Creek. This life history could be detected through PIT tag antenna arrays, if fish from neighboring streams were PIT tagged and then caught in

the Freshwater Creek out-migrant trap. These individuals should be removed before subsampling of juvenile tissue, as their inclusion in the juvenile sample would inflate the number of captures resulting in a overestimation of $\widehat{N}$.

Assumption 2: Marking in the first sampling occasion does not affect the capture probability of an individual during the second sampling occasion.

In the tGMR context an adult's reproductive success and survival to the spawning grounds controls capture probability in the second sampling occasion. Marking fish during the first sampling occasion is done via fin clip from live pre-spawning fish or carcasses of adults (pre- or post-spawning). Collecting tissue from pre-spawning adults at live traps (e.g., HFAC weir) may limit an individuals' reproductive success due to handling and impeding upstream migration, though the impacts of these factors on reproductive success is not known. In contrast, collecting tissue from post-spawn adult coho salmon will not affect reproductive success and thus this assumption is easily met when adult tissue collections are only done via carcasses.

Assumption 3: All genetic parentage assignments are correct.

Accurate estimation of escapement via tGMR requires that genetic parentage analysis produces unbiased and, ideally, error-free estimates of *R, C*', and *R*'. Parentage analysis, however, resulted in biased estimates of *R, C*', and *R*' in almost all cases (Chapter 1). Biased estimation of *R, C*' and *R*' will lead to biased estimates of escapement using tGMR. The greatest proportional bias was detected in COLONY's ability to estimate the number of parents giving rise to the juvenile subsample (*C')*. This value is the number of distinct parents detected from the genotyped juveniles in the first

sampling occasion combined with the number of distinct unsampled (inferred) parents. The number of inferred parents is constructed through half- and full-sibling reconstruction. Inferring too few parents would reduce the number of captures resulting in a negatively biased estimate of $\widehat{N}$, while inferring too many parents would increase the number of captures positively biasing the estimate of $\widehat{N}$. Bias would result from COLONY incorrectly "splitting" (Type 1 error) or "lumping" (Type 2 error) families based on sibling genotypes (Wang 2013). The frequency of Type 1 and 2 errors is dependent on number of loci used and sample sizes of both adults and juveniles (see Chapter 1, *Results*). Improvements afforded by use of more loci or multi-allelic loci are needed to eliminate bias associated with parentage analysis.

Assumption 4: All individuals have an equal probability of being captured in the first and second sampling occasion.

For this assumption to be met, capture probabilities have to be independent and equal during each of the two sampling occasions (Schwarz and Taylor 1998). In regards to reproductive success, all returning putative parents need to have an equal probability of being sampled. In Freshwater Creek, when the HFAC weir is in operation the live adult tissue collection probability is more likely to be equal because trap efficiency can be high. For carcass tissue collection within Mill and Freshwater creeks, if survey coverage is complete, or random, then this assumption should also be met by redd-count surveys. However, this assumption is likely violated because adult tissue collection is not a primary goal during many redd-count surveys, and as a result carcass collection is not a simple random sample of all returning spawners, but dependent on survey conditions. If

carcasses of any sex or size decompose, drift, or are found at equal rates then one would observe equal capture probability during the first sampling occasion. However, Pacific salmon carcass drift rates are dependent on stream flow, fish size, sex, age and the amount of instream structure (Cederholm et al. 1985; Cederholm et al. 1989; Baxter 1999; Zhou 2002). Therefore, when dealing with adult carcass sampling, one cannot assume equal capture probability during the first sampling occasion.

Using tGMR, equal capture probability during the second sampling occasion is dependent on parental reproductive success. Salmonid reproductive success varies depending on the size of the individual, timing of the return, and behavior (Dickerson et al. 2005), therefore it is likely that this assumption will be violated and the "with replacement" model will be biased. The number of recaptures is dependent on the reproductive success of the sampled adults. If the adults marked during the first sampling occasion have higher mean reproductive success then the rest of the population, this would lead to an underestimation of total escapement. If the reproductive success of the marked adults was lower than the mean of the unmarked population, this would lead to an overestimation of total escapement. In the "without replacement" model, bias may result due to heterogeneity in capture probability where the reproductive success of the marked individual drives the probability of recapturing adults marked during the first sampling occasion. This violation may be relatively minor when capture probabilities are reduced by restricting the number of offspring per spawner from many to one (Rawding et al. 2014).

Estimated Bias in Parentage Analysis for Mill and Freshwater Creek Coho Salmon Escapement Estimates

Accuracy of parentage analysis for estimation of the parameters (*R, C',* and *R'*) for both tGMR models was addressed in Chapter 1. The number of loci used, and sample sizes for both adults and juveniles for Mill and Freshwater creeks are similar to two of the simulated scenarios (*k*= 93, *M*= 40, *n*= 500 and *k*= 93, *M*= 160, *n*= 1000). For the "with replacement" model, proportional bias in *R* ranged from 1% to 5% in scenarios with similar sample sizes as the field application. This amount of bias would lead to a modest underestimation of total escapement for Mill and Freshwater creeks. For the "without replacement" model, bias is expected in both of the parameters (*C'* and *R'*) estimated by COLONY. *R'* showed positive proportional bias ranging from 1% to 3% when the adult sample was 40 and juvenile samples were 500 and 1000. The small adult sample sizes present in the Mill Creek estimates (*M*= 15, 35), would suggest modest positive proportional bias is introduced in COLONY's estimate of *R'*. This would be accompanied by proportional bias in *C*' of -40% (*k*= 93, *M*= 40 and *n*= 500). On balance, this would likely lead to a substantial underestimation of escapement. Combined, these results indicate that the tGMR estimates for Mill Creek are likely biased low. Similar patterns for Freshwater Creek likely resulted, though the adult sample size of 106 may have reduced proportional bias associated with *R'* to <1%. Also, COLONY's estimate of *C'* was likely negatively biased when sample sizes are similar to that of Freshwater Creek (*k*= 93, *M*= 160 and *n*= 1000). From the simulation results, one can expect proportional

bias of *C'* to range from -20% to -26% for the Freshwater Creek estimate. This again would lead to an underestimation of total escapement.

LITERATURE CITED

Adams, P. B., L. B. Boydstun, S. P. Gallagher, M. K. Lacy, T. McDonald, and K. E. Shaffer. 2011. California coastal salmonid population monitoring: strategy, design, and methods. California Department of Fish and Game. Fish Bulletin. 180.

Albro, P., and S. Gray. 2002. Juvenile outmigrant monitoring program 2000 report for the West Branch and East Fork of Mill Creek. Prepared for Stimson Lumber Company, Crescent City, California.

Anderson, C. W., G. Scheer, D. Ward, 2015. Results of Freshwater Creek salmonid life cycle monitoring station 2014-2015. California Department of Fish and Wildlife, Arcata, California.

Anderson, E. C., and J. C. Garza. 2005. A description of full parental genotyping. Report submitted to the Pacific Salmon Commission, Vancouver, British Columbia. 11p.

Banks, S. C., A. Horsup, A. N. Wilton, and A. C. Taylor. 2003. Genetic marker investigation of the source and impact of predation on a highly endangered species. Molecular Ecology 12(6):1663-1667.

Baumsteiger, J., and J. L. Kerby, 2009. Effectiveness of salmon carcass tissue for use in DNA extraction and amplification in conservation genetic studies. North American Journal of Fisheries Management 29(1):40-49.

Bell, E., and W. G. Duffy 2007. Previously undocumented two-year freshwater residency of juvenile coho salmon in Prairie Creek, California. Transactions of the American Fisheries Society 136(4):966-970.

Bjorkstedt, E. 2005. DARR 2.0: updated software for estimating abundance from stratified mark-recapture data. NOAA Tech. Memo. NMFS-SWSC-368. 13 p.

Bjorkstedt, E. P. 2010. DARR 2.0.2: DARR for R.
Retrieved from http://swfsc.noaa.gov/textblock.aspx?Division=FED&id=3346

Boulanger, J., and B. McLellan. 2001. Closure violation in DNA-based mark-recapture estimation of grizzly bear populations. Canadian Journal of Zoology 79(4):642-651.

Bravington, M. V., Skaug, H. J., and Anderson, E. C. 2016. Close-Kin mark recapture. Statistical Science 31(2):259-274.

Campbell, N. R., and S. R. Narum 2011. Development of 54 novel single-nucleotide polymorphism (SNP) assays for sockeye and coho salmon and assessment of available SNPs to differentiate stocks within the Columbia River. Molecular Ecology Resources 11(s1):20-30.

Cederholm, C. J., and N. P. Peterson. 1985. The retention of coho salmon (*Oncorhynchus kisutch*) carcasses by organic debris in small streams. Canadian Journal of Fisheries and Aquatic Sciences 42(6):1222-1225.

Cederholm, C. J., D.B. Houston, D. L. Cole, and W. J. Scarlett. 1989. Fate of coho salmon (*Oncorhynchus kisutch*) carcasses in spawning streams. Canadian Journal of Fisheries and Aquatic Sciences 46(8):1347-1355.

CDFG (California Department of Fish and Game). 2002. Status Review of California coho salmon north of San Francisco. Candidate Species Status Review Report 2002-3 to the California Fish and Game Commission, Sacramento, California.

Copeland, T., C. C. Kozfkay, J. Johnson, and M. R. Campbell. 2009. Do dead fish tell tales? DNA degradation in Chinook salmon (*Oncorhynchus tshawytscha*) carcasses. Northwest Science 83(2):140-147.

Crawford, B. A., and S. M. Rumsey. 2011. Guidance for monitoring recovery of Pacific Northwest salmon and steelhead listed under the federal endangered species act. National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northwest Region, Portland, Oregon.

Crisp, D. T., and P. A. Carling. 1989. Observations on siting dimensions and structure of salmonid redds. Journal of Fish Biology 34(1):119-134.

Dickerson, B. R., K. W. Brinck, M. F. Willson, P. Bentzen, and T. P. Quinn. 2005. Relative importance of salmon body size and arrival time at breeding grounds to reproductive success. Ecology 86(2):347-352.

Dreher, B. P., S. R. Winterstein, K. T. Scribner, P. M. Lukacs, D. R. Etter, G. J. Rosa, and K. B. Filcek. 2007. Noninvasive estimation of black bear abundance incorporating genotyping errors and harvested bear. The Journal of Wildlife Management 71(8):2684-2693.

Fiumera, A. C., B. A. Porter, G. D. Grossman, and J. C. Avise. 2002. Intensive genetic assessment of the mating system and reproductive success in a semi-closed

population of the mottled sculpin, Cottus bairdi. Molecular Ecology 11(11):2367-2377.

Gallagher, S. P., and C. M. Gallagher. 2005. Discrimination of Chinook salmon, coho salmon, and steelhead redds and evaluation of the use of redd data for estimating escapement in several unregulated streams in northern California. North American Journal of Fisheries Management 25(1):284-300.

Gallagher, S., P. Hahn, and D. Johnson. 2007. Redd counts. Pages 197-234. In D. Johnson, B. Shrier, J. O'Neal, J. Knutzen, X. Augerot, T. O'Neill, and T. Pearsons (editors), Salmonid field protocols hand-book: techniques for assessing status and trends in salmon and trout populations. American Fisheries Society, Bethesda, MD.

Garwood, J. M., and M. D. Larson. 2014. Reconnaissance of salmonid redd abundance and juvenile salmonid spatial structure in the Smith River with emphasis on coho salmon (Oncorhynchus kisutch). California Department of Fish and Wildlife, Arcata, CA

Hamazaki, T., and N. DeCovich. 2014. Application of the genetic mark–recapture technique for run size estimation of Yukon River Chinook salmon. North American Journal of Fisheries Management 34(2):276-286.

Hance, D. J., L. M. Ganio, K. M. Burnett, and J. L. Ebersole. 2016. Basin-scale variation in the spatial pattern of fall movement of juvenile coho salmon in the West Fork Smith River, Oregon. Transactions of the American Fisheries Society 145(5):1018-1034.

Harrison, H. B., P. Saenz-Agudelo, S. Planes, G. P. Jones, and M. L. Berumen. 2013. Relative accuracy of three common methods of parentage analysis in natural populations. Molecular Ecology 22(4):1158-1170.

Heard, W. R. 1991. Life history of pink salmon (Oncorhynchus gorbuscha). Pacific Salmon Life Histories 119-230.

Jones, A. G., and J. C. Avise. 1997. Microsatellite analysis of maternity and the mating system in the Gulf pipefish Syngnathus scovelli, a species with male pregnancy and sex-role reversal. Molecular Ecology 6(3):203-213.

Jones, O. R., and J. Wang. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources 10(3):551-555.

Justice, C. 2007. Response of juvenile salmonids to placement of large woody debris in California coastal streams. A Thesis Presented to The Faculty of Humboldt State University, Arcata, CA.

Larson, M. D. 2012. Mill Creek fisheries restoration monitoring program final report, 2010-2012 California Department of Fish and Wildlife, Arcata, CA.

Luikart, G., N. Ryman, D. A. Tallmon, M. K. Schwartz, and F. W. Allendorf. 2010. Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. Conservation Genetics 11(2):355-373.

Lukacs, P. M., and K. P. Burnham. 2005. Review of capture-recapture methods applicable to noninvasive genetic sampling. Molecular Ecology 14(13):3909-3919.

Maceina, M. J., J. Boxrucker, D. L. Buckmeier, R. S. Gangl, D. O. Lucchesi, D. A. Isermann, J. R. Jackson, and P. J. Martinez. 2007. Current status and review of freshwater fish aging procedures used by state and provincial fisheries agencies with recommendations for future directions. Fisheries 32(7):329-340.

Madej, M. A., C. O'sullivan, and N. Varnum. 1986. An evaluation of land use, hydrology, and sediment yield in the Mill Creek watershed, Northern California. Redwood National Park Research and Development Technical Report 17, Arcata, CA, 66pp.

Marucco, F., L. Boitani, D. H. Pletscher, and M. K. Schwartz. 2011. Bridging the gaps between non-invasive genetic sampling and population parameter estimation. European Journal of Wildlife Research 57(1):1–13.

McElhany, P., M. H. Ruckelshaus, M. J. Ford, T. C. Wainwright, and E. P. Bjorkstedt. 2000. Viable salmonid populations and the recovery of evolutionary significant units. NOAA Technical Memorandum NMFS-NWFSC-42.

McLeod, R. F., and C. F. Howard. 2010. Mill Creek fisheries monitoring program, final report, California Department of Fish and Game, Del Norte County, California.

Miller, L. M., M. C. Ward, and D. W. Schultz. 2015. Using genetic markers as individual tags: a case study of a mark–recapture estimate of adult muskellunge population size, North American Journal of Fisheries Management 35:2:210-215.

Morin, P. A., G. Luikart, and R. K. Wayne. 2004. SNPs in ecology, evolution and conservation. Trends in Ecology & Evolution 19(4):208-216.

Muhlfeld, C. C., M. L. Taper, D. F. Staples, and B. B. Shepard. 2006. Observer error structure in bull trout redd counts in Montana streams: Implications for inference on true redd numbers. Transactions of the American Fisheries Society 135(3):643-654.

Mulders, R., J. Boulanger, and D. Paetkau. 2007. Estimation of population size for wolverines Gulo gulo at Daring Lake, Northwest Territories, using DNA based mark-recapture methods. Wildlife Biology *13*(2):38-51.

Palsbøll, P. J., J. Allen, M. Berube, P. J. Clapham, T. P. Feddersen, P. S. Hammond, and N. Øien. 1997. Genetic tagging of humpback whales. Nature 388(6644):767-769.

Pearse, D. E., C. M. Eckerman, F. J. Janzen, and J. C. Avise. 2001. A genetic analogue of 'mark–recapture' methods for estimating population size: an approach based on molecular parentage assessments. Molecular Ecology 10(11):2711-2718.

Prentice, E. F, D. J. Maynard, S. L. Downing, D. A. Frost, M. S. Kellett, D. A. Bruland, P. Sparks-McConkey, F. W. Waknitz, R. N. Iwamoto, K. McIntyre, and N. Paasch 1994. Study to determine the biological feasibility of a new fish tagging system. Portland, Oregon, Bonneville Power Administration: 250.

Quinn, T. P. 2005. The behavior and ecology of Pacific salmon and trout. University of Washington Press, Seattle, WA.

Rawding, D.J., C.S. Sharpe, and S. M. Blankenship. 2014. Genetic-based estimates of adult Chinook salmon spawner abundance from carcass surveys and juvenile out-migrant traps. Transactions of the American Fisheries Society 143:1:55-67.

Raymond, M., and F. Rousset. 1995. An exact test for population differentiation. Evolution 49(6):1280-1283.

R Development Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rebenack, J. J., S. Ricker, C. Anderson, M. Wallace and D. M. Ward. 2015. Early emigration of juvenile coho salmon: implications for population monitoring. Transactions of the American Fisheries Society 144(1):163-172.

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

Ricker, S., and C, Anderson. 2011. Freshwater Creek salmonid life cycle monitoring station, annual report, January, 2011, California Department of Fish and Game:

Anadromous Fisheries Resource Assessment and Monitoring Program, CDFG Arcata, California.

Ricker, S.J., D. Ward, and C. W. Anderson. 2014. Results of Freshwater Creek salmonid life cycle monitoring station 2010-2013. California Department of Fish and Game, Anadromous Fisheries Resource Assessment and Monitoring Program, 50 Ericson Ct., Arcata, CA 95521.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can 191:382.

Rousset, F. 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources 8(1):103-106.

Schwarz, C. J., and C. G. Taylor, C. G. 1998. Use of the stratified-Petersen estimator in fisheries management: estimating the number of pink salmon (Oncorhynchus gorbuscha) spawners in the Fraser River. Canadian Journal of Fisheries and Aquatic Sciences 55(2):281-296.

Seamons, T. R., D. J. Rawding, P. Topping, and A. Bosworth. 2014. Spawner abundance estimates for Green River Chinook salmon. Washington Department of Fish and Wildlife, Olympia, WA.

Seamons, T. R., M. Crewson, J. Whitney, and P. Verhey. 2015. Genetic-based abundance estimates for Snohomish River Chinook salmon. Washington Department of Fish and Wildlife, Olympia, WA.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. Macmillan, New York.

Sethi, S. A., G. M. Cook, P. Lemons, and J. Wenburg. 2014. Guidelines for MSAT and SNP panels that lead to high-quality data for genetic mark–recapture studies. Canadian Journal of Zoology 92(6):515-526.

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(3):715-720.

Smith, C. T., J. Baumsteiger, W. R. Ardren, Y. Dettlaff, D. K. Hawkins, and D. M. Van Doornik. 2014. Eliminating Variation in Age at Spawning Leads to Genetic Divergence Within a Single Salmon Population. Journal of Fish and Wildlife Management 6(1):4-18.

Starks, H. A., A. J. Clemento, and J. C. Garza. 2015. Discovery and characterization of single nucleotide polymorphisms in coho salmon Oncorhynchus kisutch. Molecular Ecology Resources 16(1):277-287.

Stillwater Sciences. 2002. Mill Creek property interim management recommendations, final report. Prepared by Stillwater Sciences, Arcata, California for Save-the-Redwoods League, San Francisco and California Coastal Conservancy, Oakland, California.

Walkley. J., J. D. Deibner-Hanson, and J. M. Garwood. 2015. Mill Creek life cycle monitoring station – juvenile coho salmon out-migrant trapping project, Smith River, California. Annual Progress Report to the California Department of Fish and Wildlife Fisheries Restoration Grants Program. Smith River Alliance, California.

Wang, J. 2004. Sibship reconstruction from genetic data with typing errors. Genetics 166(4):1963-1979.

Wang, J. 2012. Computationally efficient sibship and parentage assignment from multilocus marker data. Genetics 191(1):183-194.

Wang, J. 2013. An improvement on the maximum likelihood reconstruction of pedigrees from marker data. Heredity 111(2):165-174.

Wang, J., and A. W. Santure. 2009. Parentage and sibship inference from multilocus genotype data under polygamy. Genetics 181(4):1579-1594.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38(6):1358-1370.

Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press, New York, NY.

Zhou, S. 2002. Size-dependent recovery of Chinook salmon in carcass surveys. Transactions of the American Fisheries Society 131(6):1194-1202.

SUPPLEMENTARY MATERIAL

Table 1. Summary statistics for 96 SNP loci in Freshwater and Mill Creeks. N is the number of genotyped individuals used in the analysis. H_E = expected unbiased heterozygosity, H_O = observed heterozygosity and MAF is the observed frequency of the minor allele. Asterisks (*) indicate departures from Hardy-Weinberg equilibrium. None were found to be significant ($p < 0.0001$).

		Freshwater Creek 2011-12			Freshwater Creek 2012-13			Mill Creek 2011-12			Mill Creek 2012-13		
		N=132			N=106			N=35			N=15		
Assay	References	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF
Oki94903-192	Starks et al. 2016	0.47	0.49	0.38	0.47	0.52	0.37	0.43	0.31	0.31	0.52	0.40	0.50
Oki101119-1006	Starks et al. 2016	0.06	0.06	0.03	0.01	0.01	0.00	0.45	0.37	0.33	0.48	0.47	0.37
Oki102867-667	Starks et al. 2016	0.09	0.10	0.05	0.12	0.12	0.06	0.00	0.00	0.00	0.00	0.00	0.00
Oki105115-49	Starks et al. 2016	0.19	0.20	0.11	0.22	0.21	0.12	0.47	0.37	0.37	0.51	0.60	0.43
Oki106419-292	Starks et al. 2016	0.41	0.40	0.28	0.40	0.42	0.27	0.40	0.29	0.26	0.37	0.33	0.23
Oki109874-122	Starks et al. 2016	0.11	0.11	0.06	0.27	0.28	0.16	0.29	0.34	0.17	0.37	0.47	0.23
Oki114448-101	Starks et al. 2016	0.45	0.39	0.34	0.45	0.44	0.34	0.40	0.34	0.26	0.51	0.53	0.47
Oki117815-369	Starks et al. 2016	0.50	0.57	0.50	0.38	0.36	0.25	0.18	0.20	0.10	0.24	0.27	0.13
Oki131147-353	Starks et al. 2016	0.32	0.26	0.20	0.29	0.25	0.17	0.11	0.11	0.06	0.43	0.47	0.30
Oki128757-232	Starks et al. 2016	0.39	0.36	0.27	0.39	0.43	0.27	0.45	0.37	0.33	0.43	0.33	0.30

		Freshwater Creek 2011-12			Freshwater Creek 2012-13			Mill Creek 2011-12			Mill Creek 2012-13		
		N=132			N=106			N=35			N=15		
Assay	References	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF
Oki_arp-105	Campbell and Narum 2011	0.07	0.07	0.03	0.05	0.05	0.02	0.28	0.31	0.16	0.13	0.13	0.07
Oki_p53-20	Campbell and Narum 2011	0.16	0.17	0.09	0.10	0.10	0.05	0.24	0.26	0.14	0.20	0.20	0.11
Oki109651-152	Starks et al. 2016	0.34	0.42	0.21	0.33	0.37	0.20	0.47	0.51	0.35	0.52	0.67	0.46
Oki101419-103	Starks et al. 2016	0.50	0.49	0.49	0.50	0.45	0.45	0.49	0.57	0.40	0.42	0.40	0.29
Oki103271-161	Starks et al. 2016	0.47	0.46	0.37	0.42	0.42	0.30	0.50	0.49	0.46	0.52	0.27	0.50
Oki105132-169	Starks et al. 2016	0.44	0.42	0.32	0.44	0.52	0.33	0.51	0.46	0.49	0.52	0.27	0.50
Oki106479-278	Starks et al. 2016	0.34	0.38	0.22	0.23	0.25	0.13	0.37	0.46	0.24	0.43	0.33	0.30
Oki109894-418	Starks et al. 2016	0.50	0.57	0.47	0.50	0.48	0.50	0.35	0.31	0.22	0.51	0.33	0.43
Oki114587-309	Starks et al. 2016	0.46	0.42	0.36	0.47	0.45	0.37	0.50	0.31	0.44	0.50	0.53	0.40
Oki118152-314	Starks et al. 2016	0.38	0.36	0.26	0.43	0.48	0.32	0.51	0.40	0.49	0.51	0.53	0.47
Oki123921-90	Starks et al. 2016	0.22	0.25	0.13	0.25	0.27	0.15	0.18	0.14	0.10	0.19	0.20	0.10
Oki128851-185	Starks et al. 2016	0.49	0.41	0.44	0.50	0.47	0.49	0.46	0.43	0.34	0.45	0.20	0.32
Oki110381-77	Starks et al. 2016	0.39	0.39	0.27	0.42	0.33	0.30	0.14	0.14	0.07	0.07	0.07	0.04
Oki_pigh-33	Campbell and Narum 2011	0.20	0.18	0.11	0.08	0.08	0.04	0.11	0.11	0.06	0.33	0.13	0.20

		Freshwater Creek 2011-12			Freshwater Creek 2012-13			Mill Creek 2011-12			Mill Creek 2012-13		
		N=132			N=106			N=35			N=15		
Assay	References	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF
Oki96127-66	Starks et al. 2016	0.34	0.37	0.22	0.33	0.39	0.20	0.47	0.54	0.37	0.45	0.47	0.32
Oki101554-359	Starks et al. 2016	0.03	0.03	0.02	0.14	0.13	0.08	0.06	0.06	0.03	0.07	0.07	0.03
Oki103577-70	Starks et al. 2016	0.48	0.48	0.39	0.50	0.46	0.47	0.37	0.46	0.24	0.52	0.47	0.50
Oki105235-460	Starks et al. 2016	0.36	0.30	0.23	0.40	0.39	0.28	0.41	0.46	0.29	0.51	0.60	0.43
Oki107336-45	Starks et al. 2016	0.45	0.43	0.34	0.49	0.50	0.43	0.49	0.54	0.41	0.48	0.73	0.37
Oki110064-418	Starks et al. 2016	0.03	0.03	0.02	0.04	0.03	0.02	0.11	0.11	0.06	0.07	0.07	0.03
Oki116362-411	Starks et al. 2016	0.49	0.53	0.43	0.50	0.45	0.49	0.47	0.34	0.37	0.51	0.47	0.43
Oki118175-264	Starks et al. 2016	0.28	0.23	0.16	0.16	0.11	0.08	0.43	0.54	0.31	0.33	0.27	0.20
Oki124162-62	Starks et al. 2016	0.50	0.46	0.47	0.50	0.50	0.45	0.35	0.29	0.22	0.43	0.47	0.30
Oki129870-552	Starks et al. 2016	0.13	0.14	0.07	0.13	0.14	0.07	0.39	0.29	0.26	0.40	0.27	0.27
Oki_gdh-189	Campbell and Narum 2011	0.23	0.20	0.13	0.14	0.13	0.08	0.28	0.31	0.16	0.40	0.40	0.27
Oki_rpo2j-235	Campbell and Narum 2011	0.46	0.45	0.36	0.47	0.54	0.38	0.32	0.40	0.20	0.19	0.20	0.10
Oki96158-278	Starks et al. 2016	0.50	0.53	0.46	0.50	0.41	0.49	0.39	0.34	0.26	0.00	0.00	0.00
Oki101770-525	Starks et al. 2016	0.29	0.31	0.17	0.12	0.10	0.06	0.08	0.09	0.04	0.29	0.20	0.17

		Freshwater Creek 2011-12			Freshwater Creek 2012-13			Mill Creek 2011-12			Mill Creek 2012-13		
		N=132			N=106			N=35			N=15		
Assay	References	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF
Oki103713-182	Starks et al. 2016	0.35	0.36	0.22	0.40	0.39	0.28	0.50	0.46	0.43	0.48	0.73	0.37
Oki105385-521	Starks et al. 2016	0.48	0.47	0.40	0.45	0.43	0.34	0.50	0.37	0.46	0.42	0.27	0.29
Oki107607-213	Starks et al. 2016	0.47	0.39	0.38	0.47	0.45	0.38	0.26	0.23	0.15	0.40	0.27	0.27
Oki110078-191	Starks et al. 2016	0.20	0.19	0.11	0.24	0.25	0.14	0.50	0.57	0.44	0.40	0.27	0.27
Oki116865-244	Starks et al. 2016	0.44	0.36	0.32	0.34	0.34	0.22	0.46	0.40	0.34	0.52	0.60	0.46
Oki118654-330	Starks et al. 2016	0.48	0.49	0.39	0.48	0.41	0.40	0.34	0.37	0.21	0.42	0.40	0.29
Oki125998-340	Starks et al. 2016	0.48	0.48	0.40	0.45	0.42	0.34	0.41	0.40	0.29	0.33	0.27	0.20
Oki130295-48	Starks et al. 2016	0.50	0.57	0.48	0.48	0.53	0.41	0.51	0.43	0.50	0.48	0.33	0.37
Oki109525-359	Starks et al. 2016	0.45	0.40	0.34	0.49	0.42	0.41	0.37	0.31	0.24	0.40	0.53	0.27
Oki_txnip-35	Campbell and Narum 2011	0.33	0.39	0.20	0.37	0.43	0.25	0.09	0.09	0.04	0.14	0.13	0.07
Oki96376-63	Starks et al. 2016	0.42	0.38	0.30	0.38	0.42	0.25	0.51	0.60	0.50	0.40	0.40	0.27
Oki102213-604	Starks et al. 2016	0.18	0.17	0.10	0.16	0.10	0.09	0.15	0.14	0.08	0.07	0.07	0.04
Oki104515-99	Starks et al. 2016	0.50	0.50	0.46	0.50	0.43	0.50	0.51	0.54	0.47	0.40	0.27	0.27
Oki105407-161	Starks et al. 2016	0.39	0.36	0.27	0.35	0.36	0.23	0.51	0.49	0.50	0.43	0.60	0.30
Oki107974-46	Starks et al. 2016	0.43	0.37	0.31	0.47	0.44	0.37	0.50	0.54	0.42	0.29	0.20	0.17

		Freshwater Creek 2011-12			Freshwater Creek 2012-13			Mill Creek 2011-12			Mill Creek 2012-13		
		N=132			N=106			N=35			N=15		
Assay	References	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF
Oki102267-166	Starks et al. 2016	0.43	0.44	0.31	0.38	0.25	0.25	0.39	0.43	0.26	0.00	0.00	0.00
Oki117043-374	Starks et al. 2016	0.05	0.02	0.03	0.07	0.08	0.04	0.14	0.09	0.07	0.13	0.13	0.07
Oki120024-226	Starks et al. 2016	0.44	0.52	0.33	0.46	0.42	0.35	0.25	0.23	0.14	0.33	0.27	0.20
Oki110689-43	Starks et al. 2016	0.20	0.19	0.11	0.16	0.18	0.09	0.27	0.31	0.16	0.29	0.33	0.17
Oki130524-184	Starks et al. 2016	0.43	0.45	0.31	0.43	0.39	0.32	0.47	0.54	0.36	0.43	0.60	0.30
Oki_gshpx-152	Campbell and Narum 2011	0.10	0.11	0.05	0.18	0.16	0.10	0.30	0.31	0.18	0.07	0.07	0.03
Oki_HGFA-311	Starks et al. 2016	0.44	0.41	0.32	0.47	0.36	0.38	0.36	0.03*	0.21	0.25	0.27	0.14
Oki97954-228	Starks et al. 2016	0.34	0.34	0.22	0.48	0.42	0.41	0.41	0.46	0.29	0.51	0.47	0.43
Oki102414-499	Starks et al. 2016	0.48	0.51	0.40	0.50	0.49	0.46	0.50	0.46	0.44	0.42	0.53	0.29
Oki104519-45	Starks et al. 2016	0.50	0.48	0.47	0.50	0.55	0.44	0.51	0.34	0.49	0.49	0.33	0.39
Oki105897-298	Starks et al. 2016	0.50	0.49	0.47	0.50	0.52	0.46	0.36	0.29	0.23	0.43	0.47	0.30
Oki108505-331	Starks et al. 2016	0.50	0.33	0.48	0.50	0.41	0.48	0.00	0.00	0.00	0.00	0.00	0.00
Oki111681-407	Starks et al. 2016	0.47	0.33	0.37	0.47	0.42	0.38	0.37	0.29	0.24	0.29	0.33	0.17
Oki117144-64	Starks et al. 2016	0.46	0.45	0.36	0.47	0.34	0.38	0.47	0.26	0.36	0.37	0.33	0.23
Oki120255-113-sppID	Starks et al. 2016	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

		Freshwater Creek 2011-12			Freshwater Creek 2012-13			Mill Creek 2011-12			Mill Creek 2012-13		
		N=132			N=106			N=35			N=15		
Assay	References	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF
Oki127236-383	Starks et al. 2016	0.22	0.17	0.13	0.32	0.32	0.20	0.40	0.49	0.27	0.24	0.13	0.13
Oki131460-243	Starks et al. 2016	0.28	0.32	0.17	0.33	0.31	0.21	0.50	0.46	0.46	0.52	0.73	0.50
Oki_hsc713-56	Campbell and Narum 2011	0.44	0.42	0.33	0.48	0.40	0.41	0.50	0.49	0.45	0.40	0.40	0.27
Oki_ins-167	Smith et al. 2006	0.44	0.46	0.32	0.46	0.42	0.35	0.17	0.17	0.09	0.24	0.27	0.13
Oki100771-83	Starks et al. 2016	0.31	0.32	0.19	0.19	0.21	0.10	0.11	0.11	0.06	0.07	0.07	0.03
Oki102457-67	Starks et al. 2016	0.48	0.48	0.39	0.42	0.45	0.30	0.37	0.46	0.24	0.35	0.40	0.21
Oki104569-261	Starks et al. 2016	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.09	0.04	0.19	0.20	0.10
Oki106172-60	Starks et al. 2016	0.50	0.57	0.44	0.50	0.41	0.45	0.43	0.60	0.31	0.46	0.40	0.33
Oki109243-480	Starks et al. 2016	0.05	0.05	0.03	0.06	0.06	0.03	0.18	0.20	0.10	0.19	0.20	0.10
Oki113457-324	Starks et al. 2016	0.46	0.44	0.36	0.44	0.42	0.32	0.45	0.31	0.33	0.33	0.27	0.20
Oki117286-291	Starks et al. 2016	0.23	0.22	0.13	0.26	0.26	0.15	0.06	0.06	0.03	0.07	0.07	0.03
Oki122593-430	Starks et al. 2016	0.33	0.33	0.21	0.37	0.37	0.24	0.47	0.26	0.36	0.29	0.33	0.17
Oki127760-301	Starks et al. 2016	0.04	0.03	0.02	0.12	0.11	0.07	0.31	0.26	0.19	0.37	0.20	0.23
Oki131802-368	Starks et al. 2016	0.09	0.09	0.05	0.12	0.13	0.07	0.25	0.29	0.14	0.00	0.00	0.00

		Freshwater Creek 2011-12			Freshwater Creek 2012-13			Mill Creek 2011-12			Mill Creek 2012-13		
		N=132			N=106			N=35			N=15		
Assay	References	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF
Oki_itpa-85	Campbell and Narum 2011	0.07	0.08	0.04	0.10	0.10	0.05	0.34	0.37	0.21	0.33	0.27	0.20
Oki_LWSop-554	Smith et al. 2006	0.32	0.31	0.20	0.34	0.33	0.21	0.28	0.20	0.16	0.40	0.40	0.27
Oki100974-293	Starks et al. 2016	0.29	0.31	0.18	0.34	0.36	0.22	0.08	0.09	0.04	0.40	0.53	0.27
Oki102801-511	Starks et al. 2016	0.21	0.18	0.12	0.22	0.25	0.12	0.25	0.23	0.14	0.07	0.07	0.03
Oki105105-245	Starks et al. 2016	0.15	0.14	0.08	0.24	0.23	0.14	0.08	0.09	0.04	0.00	0.00	0.00
Oki106313-353	Starks et al. 2016	0.50	0.48	0.47	0.48	0.51	0.40	0.43	0.37	0.30	0.48	0.60	0.37
Oki96222-70	Starks et al. 2016	0.39	0.43	0.26	0.36	0.35	0.23	0.47	0.49	0.37	0.51	0.33	0.43
Oki114315-360	Starks et al. 2016	0.39	0.42	0.26	0.44	0.50	0.33	0.31	0.09*	0.19	0.13	0.13	0.07
Oki117742-259	Starks et al. 2016	0.06	0.06	0.03	0.01	0.01	0.00	0.37	0.37	0.24	0.40	0.53	0.27
Oki123205-88	Starks et al. 2016	0.02	0.02	0.01	0.05	0.05	0.02	0.00	0.00	0.00	0.19	0.20	0.10
Oki128302-547	Starks et al. 2016	0.27	0.28	0.16	0.24	0.28	0.16	0.25	0.29	0.14	0.00	0.00	0.00
Oki_afp4-10	Campbell and Narum 2011	0.42	0.36	0.29	0.49	0.43	0.42	0.40	0.31	0.27	0.37	0.47	0.23
Oki_nips-159	Campbell and Narum 2011	0.37	0.43	0.24	0.39	0.40	0.26	0.50	0.37	0.44	0.33	0.40	0.20
Oki_SClkF2 R2-120	Smith et al. 2006	0.34	0.32	0.21	0.30	0.30	0.18	0.47	0.49	0.37	0.47	0.40	0.27

		Freshwater Creek 2011-12			Freshwater Creek 2012-13			Mill Creek 2011-12			Mill Creek 2012-13		
		N=132			N=106			N=35			N=15		
Assay	References	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF	H_E	H_0	MAF
	Mean	0.33	0.32	0.25	0.33	0.32	0.25	0.34	0.32	0.25	0.33	0.32	0.23
	Polymorphic Loci (%)		0.98			0.98			0.96			0.92	